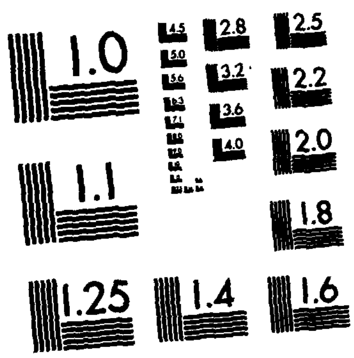
AD-A124 386 CONCEPTS AND METHODS IN MULTI-PERSON COORDINATION AND CONTROL(U) ILLINOIS UNIV AT URBANA DECISION AND CONTROL LAB T BASAR ET AL. OCT 81 DC-49 N00014-79-C-0424 1/1

UNCLASSIFIED F/G 12/1 NL

END
FILMED
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

ADA124386

COORDINATED SCIENCE LABORATORY

DECISION AND CONTROL LABORATORY

CONCEPTS AND METHODS IN MULTI-PERSON COORDINATION AND CONTROL

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

83 02 010 006

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A124 386	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) CONCEPTS AND METHODS IN MULTI-PERSON COORDINATION AND CONTROL		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER R-920(DC-49); UILU-ENG-81-2251
7. AUTHOR(s) Tamer Basar and Jose B. Cruz, Jr.		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0424 NSF ECS-79-19396 AFOSR-78-3633
9. PERFORMING ORGANIZATION NAME AND ADDRESS Coordinated Science Laboratory University of Illinois at Urbana-Champaign Urbana, Illinois 61801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Joint Services Electronics Program		12. REPORT DATE October 1981
		13. NUMBER OF PAGES 77
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release. Distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Dynamic games, Information structures, Cooperative equilibria, Noncooperative equilibrium, Stackelberg (Leader-Follower) solution concept, Nash solution concept, Stochastic team problems

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper we discuss some key concepts and methods relevant to multi-person decision-making and optimization in deterministic and stochastic dynamic systems. Specifically, we consider systems defined in discrete-time, and treat the team, Nash and Stackelberg (leader-follower) solution concepts under different information structures. We provide an up-to-date survey of the literature on these topics, and also present some new results.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

UILU-ENG-81-2251

CONCEPTS AND METHODS IN MULTI-PERSON COORDINATION AND CONTROL

by

Tamer Basar and Jose B. Cruz, Jr.

This work was supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424; in part by the National Science Foundation under Grant ECS-79-19396; and in part by the U. S. Air Force under Grant AFOSR 78-3633.

CONCEPTS AND METHODS IN MULTI-PERSON COORDINATION AND CONTROL*

TAMER BAŞAR and JOSE B. CRUZ, JR.
Department of Electrical Engineering and
Coordinated Science Laboratory
University of Illinois
1101 W. Springfield Avenue
Urbana, Illinois 61801 USA

Abstract

In this chapter we discuss some key concepts and methods relevant to multi-person decision-making and optimization in deterministic and stochastic dynamic systems. Specifically, we consider systems defined in discrete-time, and treat the team, Nash and Stackelberg (leader-follower) solution concepts under different information structures. We provide an up-to-date survey of the literature on these topics, and also present some new results.

*This paper will appear as a chapter of the book "Optimization and Control of Dynamic Operational Research Models," edited by S. G. Tzafestas, North Holland, 1982.

1. INTRODUCTION

Much of decision and control theory is concerned with a single decision-maker with a single objective function. Multiple objective functions have been considered also, but usually these are associated with a single decision-maker. Large scale systems and dynamic operations research models are likely to have a multiplicity of decision-makers. Each decision-maker may have multiple objectives. Even when each decision-maker has only one objective function, the optimization problem is significantly much more complex than that for a situation with only one decision-maker.

This chapter provides a discussion of some of the key concepts and methods that are appropriate to multiperson decision-making. When two or more decision-makers have separate objective functions, it is generally not possible to simultaneously optimize all the objective functions. One important exception is the case when all the objective functions are the same. Even in this case, the information available to each decision maker may be different from those available to others, and the problem of determining the mapping from the information space to the decision space for each decision-maker is more complex than that for a central decision-maker.

When cooperation among the decision-makers can be expected, an appropriate solution concept is that of Pareto-optimality. Otherwise, a natural concept is that of Nash equilibrium. In situations where a hierarchical decision structure is relevant, the Stackelberg or leader-follower concept is useful. These concepts will be discussed in both a deterministic and a stochastic setting.

In Section 2 we set the stage by providing motivational examples, modeling the multiperson decision problem, and defining the various solution concepts. In Section 3 we develop the concepts and methods appropriate for multiperson decision problems in deterministic systems and deterministic operations research models. The stochastic decision problem is formulated and treated in detail in Section 4. Section 5 briefly describes some examples, and Section 6 includes some concluding remarks. An extensive bibliography is included at the end of the chapter.

2. MULTI-PERSON DECISION PROBLEMS

In this section we provide a general discussion on the formulation of multi-person, and possibly multicriteria, coordination and control problems that involve uncertainty, informational decentralization and possible conflicts of interests (among the decision makers). We also discuss possible solution concepts for such decision problems. Before going into a formal presentation, let us first consider a few examples (in Section 2.1) to motivate the general formulation in the sequel.

2.1. Examples for Motivation

a) Optimum resource allocation under uncertainty

Consider a firm with (for simplicity) two divisions. The upper-level division (the headquarters) has the task of coordinating the units (of production) at the lower-level division, under incomplete information as regards to their production capabilities, availability, and quantity of resources, etc. Furthermore, there are $\underline{m}$ common resources which are to be used by some or all units in production, and therefore the headquarters has to allocate these to the units in accordance with their needs. The units may communicate their needs to the headquarters; and based on this information and some other measurements, the headquarters will have to decide on the optimal allocation that maximizes the profit to the firm (or some other appropriate utility function). One other task of the headquarters is to design an incentive scheme for remuneration of the production units, which will induce each such unit to report his true need (i.e. not to cheat in his transmittal of information) and to utilize the allocated resources most efficiently (so as,

say, to maximize the unit's share of the profit of the firm). An optimum coordination effort on the part of the headquarters will therefore force the units to behave as a team, even though the units may have their somewhat different objectives (from that of the headquarters) and operate under decentralized information.

This problem is one of multi-person coordination and control, which exhibits a hierarchy in decision making--the coordinator (headquarters) being in a position to dictate his policy on the other decision-makers (the units of the lower-level division). It also involves incomplete information, uncertainty, and a dynamic decision process with multi criteria.

b. Arms race between two nations

There is a dynamic model--known as Richardson's arms race model [117]--which describes qualitatively the armament buildup between two nations and in which the decision variables may be taken as the rates of increases or decreases in the armament levels. In making its decision as to whether to increase or decrease its current armament level, each nation will have to take a few factors into account, namely (i) the current armament level of the other nation, (ii) the economic burden associated with any possible increase in the current armament level, (iii) the response history of the other nation to past armament policies, and (iv) uncertainty associated with all this information. Yet another factor that affects the decision process is the nations' grievances and hatreds towards their "opponents". The objective of each nation will be to maximize an expected utility function that reflects a tradeoff between expected economic prosperity and national security.

This is clearly a dynamic decision process which involves two decision-makers with different objectives and whose decisions are intercoupled. It

involves uncertainty, incomplete information and noncooperative decision making.

c. Water pollution control

There are M chemical plants, located on the shores of a river, whose waste discharges pour directly into the river with no (or very little) pollution treatment. The municipality decides to take measures against this, either through a subsidy program or by penalizing those who do not properly treat their waste discharge. Assuming that the municipality is in a position to collect data from the river, the question is what type of a subsidy (or penalty) program to adopt, which will force the chemical plants to treat their waste discharges properly so that the pollution content of the river is below certain preset limits which become more stringent over the years. This is a dynamic multi-person decision problem which involves uncertainty and multi criteria. There is a conflict of interest between the municipality and the chemical plants, and there may also be some conflicts of intersts between the individual plants.

2.2. A General Formulation

A general formulation of a multi-person decision problem requires delineation of the following information:

(i) A set of decision makers (DMs), or the so called agents. Denote this set by $\bar{M} = \{1,2,..,M\}$ and a typical element by m.

(ii) An underlying probability space $(\Omega, \mathcal{B}, \mathcal{P})$ for the uncertainties, which are beyond the control of the DMs.

(iii) The length of the horizon on which the decision process is defined. Here we will adopt a discrete-time formulation with a finite horizon,

and denote the number of stages by N. Let a typical element of $\bar{N} = \{1,..,N\}$ be denoted by n.

(iv) A set of possible alternatives (<u>decisions</u>) for DM m at stage n, to be denoted by U_n^m, with $u_n^m \in U_n^m$ being a typical element. In the most general setting, U_n^m may depend on the present and past decisions of the other agents (i.e. it may not be rectangular); but here we will assume U_n^m to be rectangular for every $n \in \bar{N}$, $m \in \bar{M}$.

(v) A mathematical description of the interaction of the DMs within the system and among themselves, and with the uncertain states of the environment, i.e. specification of a system equation of the type

$$x_{n+1} = f_n(x_n, u_n^1, \ldots, u_n^M, \theta_n) \tag{2.1}$$

where $x_n, x_{n+1} \in X$ (the state space), and θ_n denotes the uncertainty affecting the outcome of the decisions at stage n. An alternate description would be specification of the probability distribution of x_{k+1} conditioned on the set of vectors $\{x_n, u_n^1, \ldots, u_n^M\}$; but we will adopt the state-space description (2.1).

(vi) An <u>information structure</u> for each DM, which characterizes the precise static or dynamic information gained and recalled by that DM at each stage of the decision process. Each such information structure will generate an appropriate information space (say Z_n^m) for DM m at stage n. In the case of deterministic information patterns, each DM will have access to some or all components of the present and past values of the state vector, as well as to the past control values of some of the other DMs. In the case of stochastic information patterns, DMs will have access to noise corrupted measurements of the state vector, say

$$y_n^m = h_n^m(x_n, \theta_n^m) \tag{2.2}$$

for DM m at stage n, where θ_n^m denotes the uncertainty corrupting the measurement. Then, the information available to DM m at stage n (denoted η_n^m) will comprise a sub-collection of the set of vectors $\{y_n^m; y_{n-1}^1, \ldots, y_{n-1}^M; \ldots; y_1^1, \ldots, y_1^M; u_{n-1}^1, \ldots, u_{n-1}^M; \ldots; u_1^1, \ldots, u_1^M\}$. If all these vectors take values in finite-dimensional spaces, then the information space Z_n^m will also be finite dimensional. [Further discussion will be devoted to this topic in the following sections; see in particular Section 4.1.]

(vii) Permissible strategies (decision laws) for each DM, defined as appropriate mappings from his information space into his decision space. Let $\pi^m = \{\gamma_1^m, \gamma_2^m, \ldots, \gamma_N^m\}$, where $\gamma_n^m : Z_n^m \to U_n^m$ is a measurable mapping. We refer to π^m as a strategy (decision law, control law) of DM m, and denote the class of all permissible strategies for DM m by Π^m. Each permissible sub-strategy γ_n^m will be assumed to belong to a sub-strategy set Γ_n^m which will have to be appropriately defined for the problem under consideration.

Permissible strategies, as introduced above, are also known as pure strategies, as opposed to mixed strategies which are defined as probability measures on $\overset{N}{\underset{n=1}{X}} \Gamma_n^m$, or behavioral strategies which are defined as independent probability measures on Γ_n^m, $n \in \bar{N}$. In the sequel we will deal only with pure strategies and refer to them simply as strategies.

(viii) An objective functional for each DM, that summarizes (mathematically) his preference ordering among different alternatives and for each fixed permissible strategy of the remaining DMs. Hence, we assume existence of a real-valued function $J^m : \Pi^1 \times \Pi^2 \times \ldots \times \Pi^M \to R$, for each $m \in \bar{M}$, which DM m strives to optimize (say minimize) by his choice of strategy $\pi^m \in \Pi^m$. Note that the effect of uncertainty (if any) is absorbed in this formulation through a possible expected utility approach. This point will be further

discussed in Section 4, where a more precise description for a subclass of problems can be found.

We should note in passing that, in the class of problems described above, the order in which each agent acts is predetermined; there exist more general formulations, however, [see, Witsenhausen (1971a)] which would allow for the order of action to be determined by a chance mechanism (which is a part of the uncertainty) and the past actions of the agents. We do not discuss such generalizations here.

2.3. Solution Concepts

The general formulation of Section 2.2 is not complete unless we specify the precise mode of decision making among the agents. Even though each agent will attempt to minimize his corresponding objective functional, this goal cannot certainly be achieved independently of the decisions of the other agents, unless the objective functional of that DM happens to be independent of all the other DMs' strategies. Hence, in order to complete the formulation of a multi-person decision problem, we have to introduce rational modes of decision making. Some selective possibilities are discussed in the sequel.

Team solution

When all agents have a common goal, we have a <u>team</u> problem with a single objective functional $J \equiv J^1 \equiv J^2 \equiv \ldots \equiv J^M$, and then an <u>optimum</u> (<u>team</u>) <u>solution</u> $\pi^* = \{\pi^{1^*}, \ldots, \pi^{M^*}\}$ is defined by

$$J(\pi^*) \leq J(\pi), \qquad \forall \pi \in \Pi \tag{2.3}$$

where we use the notation $\pi \in \Pi$ to denote $\{\pi^m \in \Pi^m,\ m \in \bar{M}\}$.

In this context, a solution concept that is somewhat weaker than the team solution is that of person-by-person optimality. Let $\pi_m \triangleq \{\pi^1,\ldots,\pi^{m-1},\pi^{m+1},\ldots,\pi^M\}$. Then, $\pi^* \in \Pi$ is person-by-person optimal if, for all $m \in \bar{M}$,

$$J^m(\pi^*) \leq J^m(\pi^*_m,\pi^m), \qquad \pi^m \in \Pi^m. \tag{2.4}$$

Note that every optimum team solution is person-by-person optimal, but not vice-versa.

Pareto-optimal solution

When the agents do not all have the same goal, but still act cooperatively, a reasonable equilibrium concept is provided by the Pareto-optimal solution. We call a subset $\Pi_p \subset \Pi$ a Pareto-optimal set if there exists no element in Π_p which is dominated by a strategy from Π, i.e. there does not exist $\pi_p \in \Pi$ and $\pi \in \Pi$ with the property

$$\begin{aligned} J^m(\pi) &\leq J^m(\pi_p) \qquad \forall m \in \bar{M} \\ \text{and} \quad J^i(\pi) &< J^i(\pi_p) \qquad \text{for at least one } i \in \bar{M} \quad . \end{aligned} \tag{2.5}$$

In other words, Π_p is the collection of nondominated strategies in Π.

Any element of the Pareto-optimal set is known as the Pareto-optimal solution for the problem under consideration, which is in general not unique. Under certain conditions [see DaCunha and Polak (1967)], the set of Pareto-optimal solutions can be obtained by considering a convex combination of the J^m's

$$J_\lambda = \sum_{m=1}^{M} \lambda_m J^m, \qquad 0 \leq \lambda_m \leq 1, \qquad \sum_{m=1}^{M} \lambda_m = 1,$$

and by minimizing $J_\lambda(\pi)$ over Π, and for fixed $\lambda = (\lambda_1, \ldots, \lambda_M)$. This yields a solution which is parameterized by λ, which generates the Pareto-optimal set.

It should be noted that a critical assumption in Pareto-optimality is cooperation. Specifically, if π^* is a Pareto-optimal solution adopted by all the agents, one of them, say the m'th one, may attain a better performance by minimizing

$$J^m(\pi_m^*, \pi^m)$$

over Π^m; but he has to refrain from adopting this policy (under the cooperative mood of decision making) since a better performance for one DM (at a Pareto solution point) necessarily implies a worse performance for some other DM.

Nash equilibrium solution

When cooperation cannot be enforced in a multi-person multi-criteria decision problem, a solution concept that safeguards against cheating by a single DM is the Nash equilibrium. We say that an M-tuple of strategies $\pi^* = \{\pi^{1*}, \ldots, \pi^{M*}\}$ provides a Nash equilibrium solution if, for all $m \in \bar{M}$,

$$J^m(\pi^*) \le J^m(\pi_m^*, \pi^m), \qquad \pi^m \in \Pi^m. \tag{2.6}$$

Note that, for the special case when J^m, $m \in \bar{M}$, are identical, this solution concept coincides with person-by-person optimality; furthermore, when $\bar{M} = \{1,2\}$, and $J^1 \equiv -J^2 \triangleq J$, we have a single inequality

$$J(\pi^{1*}, \pi^2) \le J(\pi^*) \le J(\pi^1, \pi^{2*}), \qquad \pi^1 \in \Pi^1, \quad \pi^2 \in \Pi^2 \tag{2.7}$$

which is known as the saddle-point inequality and the corresponding equilibrium solution is known as a saddle-point solution. This latter case characterizes a situation in which the two DMs have completely conflicting goals.

Stackelberg solution

Consider the class of systems with two agents and in which the roles are not symmetric. One of the DMs, known as the leader, is in a position to announce his strategy ahead of time and enforce it on the other DM, known as the follower. For each announced strategy, $\pi^1 \in \Pi^1$, of the leader, we assume that the follower acts rationally and determines his response by minimizing

$$J^2(\pi^1,\pi^2)$$

over Π^2. The set of all such solutions

$$R(\pi^1) = \{\pi^{2*} \in \Pi^2 : J^2(\pi^1,\pi^{2*}) \leq \min_{\pi^2 \in \Pi^2} J^2(\pi^1,\pi^2) \tag{2.8}$$

is known as the rational response (reaction) set of the follower. In case this is a singleton, we have the unique reaction function (mapping)

$$T : \Pi^1 \rightarrow \Pi^2, \tag{2.9a}$$

so that the leader will now determine his equilibrium strategy by minimizing

$$J^1(\pi^1,T\pi^1)$$

over Π^1. Any strategy $\pi^{1*} \in \Pi^1$ with the property

$$J^1(\pi^{1*},T\pi^{1*}) \leq J^1(\pi^1,T\pi^1), \qquad \forall \pi^1 \in \Pi^1 \tag{2.9b}$$

is known as a Stackelberg strategy for the leader. Note that T is determined here as the unique mapping satisfying

$$J^2(\pi^1,T\pi^1) \leq J^2(\pi^1,\pi^2), \qquad \forall \pi^2 \in \Pi^2 \tag{2.10}$$

for every $\pi^1 \in \Pi^1$, and with the property $T\pi^1 \in \Pi^2$. The strategy, $\pi^{2*} = T\pi^{1*}$, for the follower, that corresponds to π^{1*} under this mapping, is known as the equilibrium strategy of the follower under the Stackelberg mode of decision making.

If $R(\pi^1)$ is not a singleton, there is no unique way of defining the Stackelberg solution. One possibility is for the leader to secure his losses against nonunique rational responses of the follower, and accordingly to select a $\pi^{1^*} \in \Pi^1$ that satisfies

$$\sup_{\pi^2 \in R(\pi^{1*})} J^1(\pi^{1^*},\pi^2) \leq \sup_{\pi^2 \in R(\pi^1)} J^1(\pi^1,\pi^2), \tag{2.11}$$

for all $\pi^1 \in \Pi^1$. This, we shall also call the Stackelberg strategy for the leader.

It is also possible to extend the Stackelberg solution concept to systems with more than two DMs and possibly more than two levels of hierarchy. In this extension, if there is more than one DM at any level of hierarchy, we have to adopt either the Pareto-optimality or the Nash solution as an equilibrium concept at that particular level. As a specific case, consider an M-person decision problem with one leader and M-1 followers, and two levels of hierarchy. Suppose that there is no cooperation among the followers; then we adopt the Nash solution concept at the lower level of hierarchy and further assume that the Nash solution is unique for every $\pi^1 \in \Pi^1$. Then, there exist M-1 reaction functions $T^i : \Pi^1 \to \Pi^i$, $i=2,3,\ldots,M$, such that

$$J^i(\pi^1,\pi_i^T,T^i\pi^1) \leq J^i(\pi^1,\pi_i^T,\pi^i), \quad \pi^i \in \Pi^i, \quad i=2,3,\ldots,M, \tag{2.12a}$$

where

$$\pi_i^T = \{T^2\pi^1,T^3\pi^1,\ldots,T^{i-1}\pi^1,T^{i+1}\pi^1,\ldots,T^M\pi^1\}. \tag{2.12b}$$

A Stackelberg (hierarchical) strategy for the leader in this decision problem is a $\pi^{1^*} \in \Pi$ that satisfies

$$J^1(\pi^{1^*},T^2\pi^{1^*},\ldots,T^M\pi^{1^*}) \leq J^1(\pi^1,T^2\pi^1,\ldots,T^M\pi^1) \tag{2.13}$$

for all $\pi^1 \in \Pi^1$.

Decision problems that incorporate a hierarchy in decision making are also known as coordination problems, and the leader is referred to as the coordinator, since presence of a hierarchy enables him to coordinate the actions of the other decision makers. This is particularly true if the leader's objective function comprises a convex combination of the objective functions of the followers, in which case a Stackelberg strategy may force the followers to a Pareto-optimal solution even though they will be acting noncooperatively. Such possibilities will be discussed in the sections to follow.

3. COORDINATION AND CONTROL IN DETERMINISTIC SYSTEMS

In this section we discuss coordination and control problems in the context of deterministic systems and under deterministic information patterns. Firstly we identify deterministic problems within the framework of the formulation of §2.2 and delineate several deterministic information patterns (see §3.1). Then, we provide a brief discussion on team and Pareto-optimal solutions and representations of strategies on trajectories (in §3.2), discuss Nash equilibria (in §3.3) and Stackelberg solutions (in §3.4); finally we discuss general coordination and control problems in deterministic systems.

3.1 Deterministic Systems and Deterministic Information Patterns

The class of deterministic systems to be considered in this section will be a special case of the general formulation of §2.2, obtained by taking all probability measures to be one-point; in other words, we take the state equation to be given by

$$x_{n+1} = f_n(x_n, u_n^1, \ldots, u_n^M), \quad x_n, x_{n+1} \in R^n, \tag{3.1}$$

with the value of x_1, the initial state, specified a priori, and the stage-additive cost function to be given as

$$L^m(u^1, \ldots, u^M) = \sum_{n=1}^{N} g_n^m(x_{n+1}, u_n^1, \ldots, u_n^M, x_n) \tag{3.2}$$

for DM m.

If a decision maker has access to only the initial value of the state and does not acquire any (dynamic) information on the values of state (or controls) at other stages, we say that he has open-loop information. If, however,

he acquires perfect information concerning the current values of the state and has perfect recall on the past acquired information, we say that his information pattern is closed-loop (with memory). Hence, in the former case

$$\eta_n^m = \{x_1\}, \quad n \in \bar{N}, \tag{3.3a}$$

and in the latter case

$$\eta_n^m = \{x_n, x_{n-1}, \ldots, x_1\}, \quad n \in \bar{N}, \tag{3.3b}$$

for DM m, and these two information structures constitute the two extreme possibilities as regards deterministic information structures that involve state measurements. Two important cases "in between" are the feedback (or closed-loop no-memory) information structure in which case the decision maker recalls only the current value of the state (and also the initial state, which is known a priori), i.e.,

$$\eta_n^m = \{x_n, x_1\}, \quad n \in \bar{N}, \; m \in \bar{M}, \tag{3.4}$$

and the partial closed-loop information structure in which case the dynamic state information that the decision maker acquires and recalls is only partial, i.e.

$$\eta_n^m = \{y_n^m, y_{n-1}^m, \ldots, y_2^m, x_1\}, \; n \in \bar{N}, \; n \neq 1, \quad m \in \bar{M}, \tag{3.5a}$$

where

$$y_n^m = h_n^m(x_n), \quad n \in \bar{N}, \; n \neq 1, \quad m \in \bar{M}, \tag{3.5b}$$

and h_n^m is an appropriate function which is not necessarily one-to-one. Note

that in the partial closed-loop information structure each decision maker's current observation (or measurement) y_n^m may be different, and there is in general no sharing of information. An information structure which permits such sharing is, for example,

$$\eta_n^m = \{y_n^m, y_{n-1}, y_{n-2}, \ldots, y_2, x_1\} \qquad (3.6a)$$

where

$$y_k \triangleq \{y_k^1, y_k^2, \ldots, y_k^M\} \qquad 3.6b)$$

which is known as the <u>one-step delay observation sharing pattern</u>. It is, of course, possible to introduce other information patterns which involve sharing of only a subset of past observations and with possibly more than one stage delay. Each such information structure leads to an appropriate strategy space for each decision maker, for which we use the notation already introduced in §2.2.

3.2 Team and Pareto-optimal Solutions

When all agents have a common goal (the case of a team problem) or have different goals but act cooperatively (the case of Pareto-optimal solution), the optimum solution can be obtained by utilizing techniques of optimal control theory since in the former case there is a single objective functional to be minimized and in the latter case one may in general consider a parameterized convex combination of all the agents' cost functionals as a single objective functional to be minimized, whose parameterized solution characterizes the Pareto-optimal set. Furthermore, in order to obtain a solution under

a given general deterministic information pattern, a standard approach is first to obtain the minimizing solution under the open-loop information structure and then to synthesize a closed-loop solution as a <u>representation</u> of that open-loop solution in the strategy spaces compatible with the given dynamic information. Before discussing this point further, let us introduce the notion of "representations" of a strategy [cf. Başar (1980b)].

<u>Definition 3.1</u>. For an M-agent deterministic control (decision) problem with strategy spaces $\{\Pi^m;\ m \in \bar{M}\}$, let the strategies of all the agents, except the mth one, be fixed at $\pi^i \in \Pi^i$, $i \in \bar{M}$, $i \neq m$. Then, a strategy $\pi^m \in \Pi^m$ for DM m is a <u>representation</u> of another strategy $\tilde{\pi}^m \in \Pi^m$, with $\pi^i \in \Pi^i$ $(i \in \bar{M},\ i \neq m)$ fixed, if

(i) the M-tuples $\{\pi^m, \pi^i;\ i \in \bar{M},\ i \neq m\}$ and $\{\tilde{\pi}^m, \pi^i;\ i \in \bar{M},\ i \neq m\}$ generate the same unique state trajectory, and

(ii) π^m and $\tilde{\pi}^m$ have the same open-loop value on this trajectory. □

A salient feature of team-optimal and Pareto-optimal solutions is that under a given dynamic deterministic information structure, every representation of a solution M-tuple also constitutes a solution to the problem. However, in the cases of Nash equilibrium and Stackelberg solutions, this property no longer holds true.

3.3. <u>Nash Equilibria</u>

Derivation of Nash equilibria, when M agents have different cost functionals to minimize, involve the solution of the set of M inequalities (2.6), which, depending on the underlying information structure, may be quite a difficult problem, because each inequality defines an optimal control problem

that depends structurally on the other agents' strategies. However, if the underlying information pattern is open-loop, the structure of each of these optimal control problems is not affected by the other agents' control vectors, and hence derivation of Nash equilibria in this case becomes equivalent to solving (jointly) M optimal control problems. This argument then readily leads, by using the minimum principle, to the following set of first-order necessary conditions that yield the candidate open-loop Nash equilibrium solutions [cf. Başar (1979a)].

Proposition 3.1. For the multicriteria decision problem described by (3.1) and (3.2), let $f_n(x_n,u_n^1,..,u_n^M)$ and $g_n^m(x_{n+1},u_n^1,..,u_n^M,x_n)$ be continuously differentiable in x_n, and x_{n+1}, $n \in \bar{N}$, $m \in \bar{M}$. Then, if $\{\pi^{m^*}(x_1) = u^{m^*};\ m \in \bar{M}\}$ provides an open-loop Nash equilibrium solution and $\{x_{n+1}^*,\ n \in \bar{N}\}$ is the corresponding state trajectory, there exists a finite sequence of costate vectors $\{p_2^m,..,p_{N+1}^m\}$ for each $m \in \bar{M}$ such that the following relations are satisfied:

$$x_{n+1}^* = f_n(x_n^*,u_n^{1^*},..,u_n^{M^*}) \ ,\ x_1^* = x_1$$

$$\gamma_n^{m^*}(x_1) \equiv u_n^{m^*} = \arg\min_{u_n^m \in U_n^m} \ H_n^m(p_{n+1}^m,u_n^{1^*},..,u_n^{m-1^*},u_n^m,u_n^{m+1^*},...,u_n^{M^*},x_n^*)$$

$$p_n^m = \frac{\partial}{\partial x_n} f_n(x_n^*,u_n^{1^*},...,u_n^{M^*})'[p_{n+1}^m + [\frac{\partial}{\partial x_{n+1}} g_n^m(x_{n+1}^*,u_n^{1^*},..\mu_n^{M^*},x_n^*)]']$$
$$+ [\frac{\partial}{\partial x_n} g_n^m(x_{n+1}^*,u_n^{1^*},...,u_n^{M^*},x_n^*)]'$$

$$p_{N+1}^m = 0, \quad m \in \bar{M},\ n \in \bar{N},$$

where

$$H_n^m(p_{n+1},u_n^1,\ldots,u_n^M,x_n) \triangleq g_n^m[f_n(x_n,u_n^1,\ldots,u_n^M),u_n^1,\ldots,u_n^M,x_n)$$

$$+ p_{n+1}^{m'} f_n(x_n,u_n^1,\ldots,u_n^M); \quad n \in \bar{N},\ m \in \bar{M}. \qquad \square$$

For further discussion on the derivation of this set of first-order necessary conditions, and elucidation of some special cases as regards the structures of f_n and g_n^m, we refer to (Başar and Olsder (1982), chapter 6).

Another tractable class of problems, as far as derivation of Nash equilibria is concerned, is the class of multicriteria decision problems with closed-loop no-memory (feedback) information structure. Since every open-loop Nash equilibrium solution is also a Nash equilibrium solution under the closed-loop no-memory information structure, the Nash equilibrium solution to this class of problems cannot be unique, and in fact it exhibits "informational nonuniqueness" [see, Başar and Olsder (1982)]. One way of eliminating this informational nonuniqueness under the feedback information pattern is to require the Nash equilibrium solution to have the additional property that its restriction to the interval [n,N] is a Nash solution to the truncated version of the original problem, defined on [n,N], and this being so for all $n \in \bar{N}$. Such a solution is known as a <u>feedback Nash equilibrium solution</u>, which is free of any informational nonuniqueness, and whose derivation follows a dynamic programming type argument, as summarized in the following proposition.

<u>Proposition 3.2</u>. For the multicriteria decision problem described by (3.1) and (3.2), and under the closed-loop no-memory (or closed-loop) information pattern, the set of strategies $\{\gamma_n^{m*}(x_n);\ n \in \bar{N},\ m \in \bar{M}\}$ provides a <u>feedback Nash equilibrium</u> solution if, and only if, there exist functions $V_n^m(x)$,

$n \in \bar{N}$, $m \in \bar{M}$, such that the following recursive relations are satisfied:

$$V_n^m(x) = \min_{u_n^m \in U_n^m} \{g_n^m[\tilde{f}_n^{m^*}(x,u_n^m),\ \gamma_n^{1^*}(x),\ldots,\gamma_n^{m-1^*}(x),u_n^m,$$
$$\gamma_n^{m+1}(x),\ldots,\gamma_n^{M^*}(x),x] + V_{n+1}^m[\tilde{f}_n^{m^*}(x,u_n^M)]\}$$
$$= g_n^m[\tilde{f}_n^{m^*}(x,\gamma_n^{m^*}(x)),\gamma_n^{1^*}(x),\ldots,\gamma_n^{M^*}(x),x] + V_{n+1}^m[\tilde{f}_n^{m^*}(x,\gamma_n^{m^*}(x))]$$

$$V_{N+1}^m(x) = 0,\quad m \in \bar{M}$$

where

$$\tilde{f}_n^{m^*}(x,u_n^m) \triangleq f_n[x,\gamma_n^{1^*}(x),\ldots,\gamma_n^{m-1^*}(x),u_n^m,\gamma_n^{m+1^*}(x),\ldots,\gamma_n^{M^*}(x)]. \qquad \Box$$

It should be clear from the above that feedback Nash equilibrium solution can be obtained recursively, by solving a set of static Nash problems at each stage, which is a feature that makes it computationally attractive. Yet another important feature that should be recorded is that feedback Nash solution is indeed a Nash equilibrium solution under the closed-loop no-memory or closed-loop information patterns (satisfying inequalities (2.6)), but one of many "informationally nonunique" equilibria under those dynamic information structures.

As already mentioned above, when we have the closed-loop information pattern, or any dynamic information pattern that exhibits redundancy in information, Nash equilibria are informationally nonunique and there exists in fact an uncountable number of such equilibria. A set of reasons for this is now provided in the following definition and proposition, where a proof for the latter can be found (Başar and Olsder (1982), chapter 6).

Definition 3.2. Let A and B be two M-person N-stage deterministic multicriteria decision problems which admit precisely the same extensive form description (as in §2.2) except the underlying information pattern (and, of course, also the strategy spaces whose descriptions depend on the information pattern). Let η_A^m (respectively, η_B^m) denote the information pattern of DM m in problem A (respectively, B), and let the inclusion relation $\eta_A^m \subseteq \eta_B^m$ imply that whatever DM m knows at each stage of A he also knows at the corresponding stages of B, but not necessarily vice versa. Then, A is informationally inferior to B if $\eta_A^m \subseteq \eta_B^m \ \forall\ m \in \bar{M}$, with strict inclusion for at least one m. □

Proposition 3.3. Let A and B be two deterministic decision problems as introduced in Definition 3.2, so that A is informationally inferior to B. Furthermore, let the strategy spaces of the decision makers in the two problems be compatible with the given information patterns and constraints (if any) imposed on the controls, so that $\eta_A^m \subseteq \eta_B^m$ implies $\Pi_A^m \subseteq \Pi_B^m$, $m \in \bar{M}$. Then, (i) any Nash equilibrium solution for A is also a Nash equilibrium solution for B, (ii) if $\{\pi^1,..,\pi^M\}$ is a Nash equilibrium solution for B such that $\pi \in \Pi_A^m$ for all $m \in \bar{M}$, it is also a Nash equilibrium solution for A. □

Hence, multicriteria deterministic decision problems with dynamic information patterns that exhibit redundancy in information are not well defined under the Nash solution concept (since they admit a plethora of informationally nonunique equilibria) unless some additional selection criteria are introduced --such as the requirements imposed by the feedback Nash solution discussed earlier. We do not pursue this point any further here, but note that one such criterion is in fact provided in §4.3 under a stochastic set-up.

3.4. Stackelberg (Leader-Follower) Solutions

In this subsection, we treat the problem of optimal control and coordination of deterministic systems under a hierarchical decision structure, and investigate derivation of optimal control and coordination strategies by employing the Stackelberg solution concept introduced in §2.3. As discussed earlier in §2.3, while introducing the Stackelberg solution concept, existence of a hierarchy in decision making results in an asymmetry in the roles of the agents, with some of them being in a position to dictate their strategies on the others.

In general, derivation of Stackelberg solutions in dynamic decision problems is quite challenging, the difficulty being mostly of conceptual nature. However, for some special information structures, the problem becomes tractable because some standard methods and techniques of optimization and optimal control theory become applicable. One such class of problems is characterized by open-loop information structure, and say two agents (i.e. $M = 2$) for the sake of simplicity in the discussion to follow. Since the leader's information structure is open-loop, the optimization problem faced by the follower in the determination of his optimal response set (2.8) is structurally independent of different choices of strategies by the leader, and therefore the first phase of the derivation of the Stackelberg solution is a feasible (tractable) optimal control problem. In particular, if the follower's cost functional is strictly convex in his control, the rational response set $R(\pi^1)$ will be a singleton and the reaction function T [see (2.9a)] will be determined completely by a set of necessary and sufficient conditions which, under certain structural assumptions on f_n and g_n^2, $n \in \bar{N}$, will lead to an analytical solution for T. If such an

analytic solution can be found, then the leader's optimization problem $\min_{\Pi^1} J^1(\pi^1, T\pi^1)$ is again a standard optimal control problem which can readily be solved using the available techniques for dynamic optimization, and the open-loop representation of this solution (in case it is obtained as closed-loop solution) will constitute a Stackelberg strategy for the leader. In case an analytic expression for T does not exist, the necessary and sufficient conditions that describe T will have to be treated as constraints in the leader's optimization problem which again involves no difficulties of conceptual nature. A set of equations from which the solution of this constrained optimal control problem can be obtained can be found in (Başar and Olsder (1982), chapter 7); we do not discuss this class of problems any further here. It is worth noting here that the preceding derivation is valid not only under the open-loop information structure for both agents, but also when the follower has access to dynamic state information --the only requirement is that the leader should have only open-loop information. Furthermore, one can envisage direct extensions of this procedure to M-agent problems with one leader and M-1 followers, with the latter determining their policies according to the Nash or Pareto-optimum solution conept, and with the leader having access to only open-loop information; there appears to be no difficulties of conceptual nature in such an extension.

When the leader has access to dynamic state information, derivation of the Stackelberg solution constitutes a challenging problem, and the standard techniques of optimization do not apply, since the optimal control problem characterizing the rational response set $R(\pi^1)$ is now structurally dependent on the leader's choice or strategies. One way out of this difficulty

would be to fix the structure of the leader's possible strategies parametrically, find the follower's rational response as a function of these parameters and then optimize the leader's cost functional over these parameter values, also in view of the follower's response; this definitely leads to suboptimal strategies for the leader --the degree of suboptimality depending on how representative the fixed structure is in the general class of policies.

Another way of making the Stackelberg problem tractable is to require the solution have a feedback property (under the closed-loop no-memory of closed-loop information sharing pattern), analogous to the case of the feedback Nash equilibrium solution, which would lead to a recursive derivation in retrograde time that involves solution of static Stackelberg problems at every stage. The solution obtained through such a recursive procedure is called a <u>feedback Stackelberg solution</u> [cf. Simaan and Cruz (1973a,b)] and satisfies the conditions given in the following proposition.

<u>Proposition 3.4</u>. For the two-agent multicriteria decision problem described by (3.1) and (3.2) with $M = 2$, and under the closed-loop no-memory (or closed-loop) information structure, the set of strategies $\{\gamma_n^{1^*}(x_n), \gamma_n^{2^*}(x_n);\ n \in \bar{N}\}$ provides a <u>feedback Stackelberg</u> solution with DM 1 as leader, if

$$\min_{\gamma_n^1 \in \Gamma_n^1,\, \gamma_n^2 \in R_n(\gamma_n^1)} \tilde{G}_n^1(\gamma_n^1, \gamma_n^2, x_n) = \tilde{G}_n^1(\gamma_n^{1^*}, \gamma_n^{2^*}, x_n), \text{ for all } x_n \in X,\ n \in \bar{N},$$

where $R_n(\gamma_n^1)$ is a <u>singleton</u> set defined by

$$R_n(\gamma_n^1) = \{\beta_n^2 \in \Gamma_n^2 : \tilde{G}_n^2(\gamma_n^1, \beta_n^2, x_n) = \min_{\gamma_n^2} \tilde{G}_n^2(\gamma_n^1, \gamma_n^2, x_n)\},$$

$$\tilde{G}_n^m(\gamma_n^1,\gamma_n^2,x_n) \triangleq G_n^m[f_n(x_n,\gamma_n^1(x_n),\gamma_n^2(x_n)),\gamma_n^1(x_n),\gamma_n^2(x_n),x_n),$$

$$m = 1,2, \quad n \in \bar{N},$$

and G_n^m is defined recursively by

$$G_n^m[x_{n+1},\gamma_n^1(x_n),\gamma_n^2(x_n),x_n] = G_{n+1}^m[f_{n+1}(x_{n+1},\gamma_{n+1}^{1^*}(x_{n+1}),\gamma_{n+1}^{2^*}(x_{n+1})),$$

$$\gamma_{n+1}^{1^*}(x_{n+1}),\gamma_{n+1}^{2^*}(x_{n+1}),x_{n+1}] + g_n^m \; ; \quad G_{N+1}^m \equiv 0, \; m = 1,2. \quad \square$$

The feedback Stackelberg solution corresponds to the case when the leader can enforce his strategy on the follower only stagewise; however, if he has the power and ability to declare and enforce his strategy several stages in advance throughout the decision process, or from the very beginning for the entire duration of the decision process, the cost that the leader incurs will definitely be less (or at least not higher) than his optimal cost under the feedback Stackelberg solution. In other words, in contrast to the feature recorded after Proposition 3.2 in the case of feedback Nash solution, the feedback Stackelberg solution is <u>not</u> necessarily a Stackelberg solution, i.e. it need not satisfy (2.9b); conversely, a Stackelberg solution obtained under the closed-loop no-memory information structure is not necessarily a feedback Stackelberg solution. On the other hand, derivation of a Stackelberg solution under dynamic state equation is a relatively much more difficult problem, for which the standard techniques of optimization cannot be used.

Another case treated in the literature recently is the closed-loop

no-memory information structure where the leader's strategy is a function of the current state. This problem leads to a nonclassical control problem where the partial derivative of the leader's strategy with respect to the state appears. It is shown in Papavassilopoulos and Cruz (1979a) that the optimal values of the state, controls, and objective functions are not changed by using controls which are more general than affine functions of the state. When the measurement is a function of the state (possibly nonlinear) the strategy may be assumed to be affine in the measurement without loss of generality.

Quite recently, an indirect approach has been developed towards the solution of such nonclassical optimization-decision problems when the leader has access to redundant information (such as the closed loop state information). In the sequel we discuss some aspects of this new approach and derivation of the dynamic Stackelberg solution.

Now, for the general two-agent decision problem of this subsection, and with the leader having access to closed-loop state information, consider the following sequence of optimization problems.

STEP 1. For a fixed set of state vectors $\{x_n,\ n \in \bar{N},\ n \neq 1\}$ say $\{x_n = \bar{x}_n,\ n \in \bar{N},\ n \neq 1\}$, and leader's control vectors $\{u_n^1,\ n \in \bar{N}\}$, minimize

$$g_N^2(x_{N+1}, u_N^1, u_N^2, \bar{x}_N) + \sum_{n=2}^{N-1} g_n^2(\bar{x}_{n+1}, u_n^1, u_n^2, \bar{x}_n) + g_o^2(\bar{x}_2, u_1^1, u_1^2, x_1) \qquad (3.7)$$

over $u_n^2 \in U_n^2$, $n \in \bar{N}$, and subject to the constraint

$$\begin{aligned} x_{N+1} &= f_N(\bar{x}_N, u_N^1, u_N^2) \\ \bar{x}_{n+1} &= f_n(\bar{x}_n, u_n^1, u_n^2), \quad n \leq N-1, \quad \bar{x}_1 = x_1. \end{aligned} \qquad (3.8)$$

Denote the solution of this problem by

$$u_n^2 = z_n(\bar{x}_2,..,\bar{x}_N;u_1^1,..,u_N^1), \quad n \in \bar{N}. \tag{3.9}$$

STEP 2: Now consider minimization of the function

$$g_N^1(x_{N+1},u_N^1,u_N^2,\bar{x}_N) + \sum_{n=2}^{N-1} g_n^1(\bar{x}_{n+1},u_n^1,u_n^2,\bar{x}_n) + g_0^1(\bar{x}_2,u_1^1,u_1^2,x_1) \tag{3.10}$$

over the leader's controls $\{u_n^1 \in U_n^1,\ n \in \bar{N}\}$, and the state values $\{\bar{x}_n,\ n \in \bar{N},\ n \neq 1\}$, subject to (3.8) and (3.9). Denote the minimizing solution by $\{u^{1^*},\ n \in \bar{N}\}$ and $\{x_n^*,\ n \in \bar{N},\ n \neq 1\}$ and the corresponding value of expression (3.10) by J^{1^*}.

The quantity J^{1^*}, thus obtained, provides, under a fairly general set of conditions, a tight lower bound on the Stackelberg cost $J^1(\pi^{1^*},T\pi^{1^*})$ of the leader (as defined by 2.9b)). These conditions basically involve existence of a strategy $\pi^{1^*} \in \Pi^1$, for the leader, which is

(i) a closed-loop representation of the open-loop policy $\{u_n^{1^*}, n \in \bar{N}\}$ on the trajectory $\{x_n = x_n^*,\ n \in \bar{N}\}$, where $u_n^{1^*}$ and x_n^* are as defined above, with $x_1^* = x_1$;

(ii) forces the minimum value of (3.7) to be attained at $\{u_n^2 = z_n(x_2^*, ..,x_N^*;u_1^{1^*},..u_N^{1^*}),\ n \in \bar{N}\}$, with the minimization problem defined by replacing u_n^1 in (3.7) and (3.8) by $\gamma_n^{1^*}(.)$, $n \in \bar{N}$, and $\bar{x}_n$, $n \in \bar{N}$, in (3.8) by x_n, and retaining this new form of (3.8) as a constraint. Note that this latter requirement is equivalent to the statement that the follower's rational response to the leader's announced strategy π^{1^*} should lead to the trajectory $\{x_n^*, n \in \bar{N}\}$ and have the open-loop representation $\{u_n^2 = z_n(x_2^*,..,x_N^*;u_1^{1^*},..,u_N^{1^*}),$

$n \in \bar{N}\}$.

Several recent papers have investigated, in special contexts, satisfiability of these two conditions, and derivation of corresponding strategies (π^{1*}) for the leader. Başar and Selbuz (1979a,b) have shown that when the system equation is linear and cost functionals are quadratic, there are cases when J^{1*} coincides with the global minimum value of J^1 (in particular, if the follower does not act in the last stage of the game) and a corresponding Stackelberg strategy for the leader is of the linear, one-step memory type. Tolwinski (1981) has shown that for the same class of problems, use of nonlinear strategies by the leader extends the parameter region for which the preceding properties of the solution hold true. Papavassilopoulos and Cruz (1980), Başar and Olsder (1980) and Başar (1981d) have investigated counterparts of these results and their extensions in the continuous time. Ho, Luh and Muralidharan (1980), Ho, Luh and Olsder (1980), and Salman and Cruz (1981) have drawn parallels between these results and incentive scheme design problems in economics and have discussed applications of these concepts to microeconomics and social choice theory. Başar (1981e) and Tolwinski (1980) have discussed possible extensions to multi-agent cases when there exist more than two levels of hierarchy and several agents at every level of decision making. Başar and Selbuz (1979b) show that if there exist two levels of hierarchy and more than one agent in the follower's group, the leader can still retain his powerful position by announcing an appropriate linear one-step memory strategy (for linear-quadratic problems) that would force the followers (who are making their decisions noncooperatively and under the Nash solution concept) to minimize globally the leader's cost function. Başar (1980b) has further discussed coordination aspects of such problems and has investigated the possibilities for the

leader to coordinate the followers in such a way that the resulting solution will be Pareto-optimum, even though the followers may be acting noncooperatively.

It is possible to extend the two-step derivation of the closed-loop Stackelberg solution, outlined earlier and defined through the optimization problems (3.7) - (3.10), to the case when the leader's information is partial closed-loop [see (3.5)]. In this case the two optimization problems at Steps 1 and 2 will be replaced, respectively, by the following:

<u>STEP 1'</u>: Let the observation vector y_n^1, defined by (3.5b), belong to the space Y_n^1. For a fixed set of observation vectors $\{y_n^1 \in Y_n^1,\ n \in \bar{N},\ n \neq 1\}$,say $\{y_n^1 = \bar{y}_n,\ n \in \bar{N},\ n \neq 1\}$, and leader's control vectors $\{u_n^1,\ n \in \bar{N}\}$, minimize

$$\sum_{n=0}^{N} g_n^2(x_{n+1},u_n^1,u_n^2,x_n) \tag{3.11}$$

over $u_n^2 \in U_n^2$, $n \in \bar{N}$, and subject to the constraints

$$x_{n+1} = f_n(x_n,u_n^1,u_n^2) \tag{3.12a}$$

$$h_n^1(x_n) = \bar{y}_n, \quad n \in \bar{N},\ n \neq 1. \tag{3.12b}$$

Denote the solution of this optimization problem by

$$u_n^2 = z_n(\bar{y}_n,\ldots,\bar{y}_N;u_1^1,\ldots u_N^1). \tag{3.13}$$

<u>STEP 2'</u>: Now minimize the function

$$\sum_{n=1}^{N} g_n^1(x_{n+1},u_n^1,u_n^2,x_n) \tag{3.14}$$

over the leader's controls $\{u_n^1 \in U_n^1,\ n \in \bar{N}\}$, and the measurement values $\{y_n^1 \in Y_n^1,\ n \in \bar{N},\ n \neq 1\}$, subject to (3.12a), (3.13) and

$$y_n^1 = h_n^1(x_n),\ n \in \bar{N},\ n \neq 1.$$

Denote the minimizing solution by $\{u_n^{1^*},\ n \in \bar{N}\}$ and $\{y_n^*,\ n \in \bar{N},\ n \neq 1\}$, the resulting state trajectory by $\{x_n^*,\ n \in \bar{N}\}$ and the corresponding values of expression (3.14) by J^{1^*}.

The conditions for J^{1^*} to provide a tight lower bound on the Stackelberg cost $J^1(\pi^{1^*}, T\pi^{1^*})$ involve, in this case, existence of a strategy $\pi^{1^*} \in \Pi^1$ [Π^1 is defined here as the class of all mappings compatible with the information structure η_n^1 given by (3.5a)] that satisfy condition (i) in the perfect information case and, in addition

(ii′) forces the minimum value of (3.11) to be attained at $\{u_n^2 = z_n(y_2^*, \ldots, y_N^*;\ u_1^{1^*}, \ldots, u_N^{1^*}),\ n \in \bar{N}\}$ with the minimization problem defined by replacing u_n^1 in (3.11) and (3.12a) by $\gamma_n^{1^*}(.)$, $n \in \bar{N}$, by replacing $\bar{y}_n$ in (3.12b) by y_n^1, and by retaining these new forms of (3.12a)-(3.12b) as constraints.

For further details on the satisfiability of these two conditions and derivation of dynamic Stackelberg solution under partial state information, we refer to Başar (1980c) and Zheng and Başar (1981); the latter reference also investigates existence and derivation of affine Stackelberg strategies in such problems.

4. COORDINATION AND CONTROL IN STOCHASTIC SYSTEMS

In this section we discuss coordination and control problems in the context of stochastic systems and under both deterministic and stochastic information patterns. We first delineate (in § 4.1) several different information structures that we shall encounter in our analysis, and then discuss (in § 4.2) derivation and properties of optimal solutions in stochastic team problems. Subsequently in §4.3 we discuss Nash equilibria and in §4.4 the Stackelberg solution, for stochastic systems and under different information patterns.

4.1. Information Structures in Stochastic Systems

In stochastic systems we encounter two general classes of information patterns, viz. deterministic and stochastic patterns:

a) Deterministic information structures

We have discussed these thoroughly in § 4.1 in the context of deterministic systems. The same patterns, namely, closed-loop perfect state, feedback, one-step (k-step) delay perfect state, and partial closed-loop information structures, are appropriate also in stochastic systems, whenever the agents have access to the value of the initial state and to some deterministic information on the current and/or past values of the state.

b) Stochastic information structures

Assume that each agent has access to noisy measurement on the current value of the state through a measurement equation of the type (2.2), and that agents are also in a position to exchange some of their information (with or without delay). In such a case we have basically three general types of information structures as described below:

i) <u>Centralized information pattern</u>: All agents exchange their measurements without any delay, and also recall their past information, i.e.

$$\eta_n^m = \{y_n, y_{n-1}, \ldots, y_1\}, \; m \in \bar{M}, \; n \in \bar{N} \tag{4.1}$$

where

$$y_k \triangleq (y_k^1, y_k^2, \ldots, y_k^M), \; k \in \bar{N}.$$

This is also known as a <u>classical</u> information pattern, and it could also involve the past control laws, i.e.

$$\eta_n^m = \{y_n, y_{n-1}, \ldots, y_1; u_{n-1}, u_{n-2}, \ldots, u_1\}, m \in \bar{M}, \; n \in \bar{N} \tag{4.2}$$

where

$$u_k \triangleq (u_k^1, u_k^2, \ldots, u_k^M).$$

The two information structures (4.1) and (4.2) are <u>not</u> equivalent (even though they generate the same sigma-field for each fixed set of control laws), but only in team problems may they be used interchangeably without affecting the minimum value of the common objective functional --a point which will be further discussed in § 4.2.

ii) <u>Quasi-classical information patterns</u>: In this group we have the "one-step delay observation (measurement) sharing pattern", in which case

$$\eta_n^m = \{y_n^m, y_{n-1}, \ldots, y_1\}, \; m \in \bar{M}, \; n \in \bar{N}, \tag{4.3a}$$

and the "one-step delay information sharing pattern" with

$$\eta_n^m = \{y_n^m, y_{n-1}, \ldots, y_1;\ u_{n-1}, u_{n-2}, \ldots, u_1\},\ m \in \overline{M},\ n \in \overline{N}. \quad (4.3b)$$

In the former case all measurements are shared with a delay of one stage, while in the latter case also the past control values are shared. Our earlier comments regarding the equivalence of (4.1) and (4.2) are equally valid here in the context of (4.3a) and (4.3b); more discussion on this issue will be included in § 4.2.

A more general type of a quasi-classical information structure is the so-called <u>partially nested</u> information structure which we introduce next. Towards this end, assume that the joint probability distribution of the random variables associated with the stochastic system (2.1) and the measurement system (2.2) is independent of the values of the state and the controls. Then, by iterative substitution, (2.1) can be written as

$$\begin{aligned} x_{n+1} &= f_n(x_n, u_n, \theta_n) \\ &= f_n[f_{n-1}(x_{n-1}, u_{n-1}, \theta_{n-1}), u_n, \theta_n] \triangleq f_n^{n-1}(x_{n-1}; u_n, u_{n-1}; \theta_n, \theta_{n-1}) \\ &\quad \text{---} \\ &= f_n^1(x_1; u_n, u_{n-1}, \ldots, u_1; \theta_n, \theta_{n-1}, \ldots, \theta_1), \end{aligned} \quad (4.4a)$$

and thus the state at any stage can be expressed solely in terms of the past controls, the past noise vectors and the initial state. In terms of this notation, the measurement equation (2.2) can be written as

$$\begin{aligned} y_n^m &= h_n^m[f_{n-1}^1(x_1;\ u_n, u_{n-1}, \ldots, u_1;\ \theta_n, \ldots, \theta_1), \theta_n^m] \\ &\triangleq H_n^m(x_1;\ u_n, \ldots, u_1;\ \theta_n, \ldots, \theta_1;\ \theta_n^m); \end{aligned} \quad (4.4b)$$

that is, in this new form it depends only on the "primitive" random variables and the control vectors. Now, we call an information structure $\{\eta_n^m \subseteq \{y_n^1,..,y_n^M;$ $y_{n-1}^1,..,y_{n-1}^M;...;\ y_1^1,..,y_1^M;\ u_{n-1},..,u_1\},\ n \in \overline{N},\ m \in \overline{M}$ <u>partially nested</u> if whenever η_n^m depends on u_k^i for some $k \leq n$ and $i \in \overline{M}$ [either directly or through a measurement equation in the form (4.4)], the inclusion relation $\eta_n^m \supseteq \eta_k^i$ holds[†] --this being so for every such dependence. In other words, if an information structure is partially nested, an agent's information at a particular stage n can depend on the control of some other agent at some stage $k \leq n$ only if he also has access to the information available to that agent at that stage k.

The one-step-delay observation sharing pattern and the one-step delay information sharing pattern introduced earlier are special types of partially nested information patterns. The reason why we are interested in partially nested information patterns is that stochastic optimization and in particular team problems with such information patterns are considerably more tractable than those with nonclassical information patterns --this latter concept to be defined in the sequel.

iii) <u>Nonclassical information patterns</u>. An information pattern is said to be <u>nonclassical</u> if it is not partially nested. Equivalently, if $\{\eta_n^m,$ $n \in \overline{N},\ m \in \overline{M}\}$ is nonclassical, there exists some set of indices $\{n,\ k \in \overline{N},$ $m,\ i \in \overline{M},\ n \geq k\}$ such that η_n^m depends on u_k^i but $\eta_n^m \not\supseteq \eta_k^i$.

[†] This inclusion relation can be replaced by the somewhat more general requirement that "the elements of η_k^i can be recovered by measurable transformations on the elements of η_n^m".

4.2. Solutions of Stochastic Team Problems

Under the centralized stochastic or deterministic information patterns, stochastic team problems become equivalent to stochastic control problems, and therefore the solution techniques developed for this latter class of problems [see e.g. Bertsekas (1976)] are directly applicable to team problems. In this context, it is immaterial whether the agents also have access to values of past controls, since there is a single goal and a single objective functional to minimize. In other words, the minimum value of a team cost functional J will be the same regardless of whether it is computed under (4.1) or (4.2); in that sense we call the two information structures equivalent as far as the optimal team solution is concerned. However, this feature is no longer valid in multi-criteria problems (under Nash or Stackelberg solution concepts).

If the underlying information structure is not centralized, the derivation of the optimal team solution is in general quite intractable. For some special types of stochastic team problems and under the partially nested information pattern, however, the derivation becomes tractable by conversion into an equivalent static formulation. Before discussing this conversion, we first state a related result [Proposition 4.1] on an important property of partially nested information patterns in stochastic team problems:

Let $\{\eta_n^m \subseteq \{y_n^1,\ldots,y_n^M;\ y_{n-1}^1,\ldots,y_{n-1}^M;\ldots;\ y_1^1,\ldots,y_1^M;\ u_{n-1},\ldots,u_1\}$, $n \in \overline{N}$, $m \in \overline{M}\}$ be a partially nested information pattern, with the corresponding strategy spaces denoted by $\{\Pi^m, m \in \overline{M}\}$ and the corresponding sub-strategy spaces by $\{\Gamma_n^m, n \in \overline{N}, m \in \overline{M}\}$. Let $\hat{\eta}_n^m$ denote, for each $n \in \overline{N}$, $m \in \overline{M}$, the inter-

section of the finite sets η_n^m and $\{y_n^1,..,y_n^M;\ y_{n-1}^1,..,y_{n-1}^M;...;\ y_1^1,..,y_1^M\}$. Note that $\{\hat{\eta}_n^m,\ n \in \overline{N},\ m \in \overline{M}\}$ is also a partially nested information structure, which does not involve any explicit dependence on past control vectors [whereas η_n^m may explicitly depend on controls]. Denote the corresponding strategy spaces by $\{\hat{\Pi}^m,\ m \in \overline{M}\}$ and the sub-strategy spaces by $\{\hat{\Gamma}_n^m,\ n \in \overline{N},\ m \in \overline{M}\}$. Consider a stochastic team problem with cost functional $J(\pi^1,\pi^2,..,\pi^M)$ to be jointly minimized [over $\overset{M}{\underset{m=1}{X}} \Pi^m$] by all agents. Then, we have the following important result.

<u>Proposition 4.1.</u>

(i) To every fixed M-tuple $(\gamma_1^m,\ \gamma_2^m,..,\gamma_N^m) \triangleq \pi^m \in \Pi^m,\ m \in \overline{M}$, there corresponds a unique set of strategies $\{\hat{\pi}^m = (\hat{\gamma}_1^m,\hat{\gamma}_2^m,..,\hat{\gamma}_N^m),\ m \in \overline{M}\}$ such that the sigma-field generated by η_n^m with $u_n^m = \gamma_n^m(\eta_n^m)$, $n \in \overline{N}$, $m \in \overline{M}$, is <u>equivalent</u> to the sigma-field generated by $\hat{\eta}_n^m$ with $u_n^m = \hat{\gamma}_n^m(\hat{\eta}_n^m)$, $n \in \overline{N}$, $m \in \overline{M}$.

(ii) J admits a global minimum over $\overset{M}{\underset{m=1}{X}} \Pi^m$, if and only if it admits a global minimum over $\overset{M}{\underset{m=1}{X}} \hat{\Pi}^m$, and the minimum values of J in both cases are the same. □

This proposition is a consequence of the ob ervation that, under the partially nested information structure, any direct information concerning the value of control is redundant since it can be recovered from the measurement information once the control law is known. Consequently, additional information concerning the values of past controls [provided that we still have a partially nested information structure] does not help to improve upon the globally optimal team solution. An implication of this property is that, given a specific partially nested information pattern for a stochastic team problem, we can construct an equivalent (larger or smaller) partially nested

information structure that is equivalent to it, and reconsider the original team problem under this new information structure without affecting the optimal value of the objective functional. What we gain in return for such a conversion is a possible simplification in the derivation of the optimal team solution. The team solution obtained under the new information structure can then be expressed in terms of the original information structure. Examples of such an indirect derivation of optimal solutions in stochastic team problems can be found in Ho and Chu (1972) and Bagchi and Başar (1980), and they are primarily linear quadratic problems. A CAVEAT for the reader, at this point, is that neither Proposition 4.1 nor any of these conversion techniques have counterparts in multi-criteria problems (under Nash or Stackelberg solution concept).

Let us now consider one special class of stochastic team problems in some detail. Assume that the information structure is partially nested, and that the measurement equations (4.4b) are separable in the control variables, i.e. (4.4b) can be written as

$$y_n^m = \tilde{H}_n^m(x_1;\ \theta_n,\ldots,\theta_1;\ \theta_n^m) + \tilde{G}_n^m(u_n,\ldots,u_1),\ n \in \overline{N},\ m \in \overline{M}. \tag{4.5}$$

Here, the function $\tilde{G}_n^m$ depends on the control vectors in a way that is consistent with the underlying partially nested information structure $\{\eta_n^m;\ n \in \overline{N}, m \in \overline{M}\}$; i.e. $\tilde{G}_n^m$ is a function of u_k^i only if η_n^m includes η_k^i. Now, if $\{\pi^{m*} \in \Pi^m, m \in \overline{M}\}$ denotes an optimal team solution for a stochastic team problem with such a partially nested information structure, and with a cost functional J, where

$$J(\pi) = E\{L(\xi,u^1,\ldots,u^M) \mid u^m = \pi^m(\eta^m),\ m \in \overline{M}\ ,$$

and ξ denotes the collection of all primitive random variables, we have (from the definition of team optimality)

$$J(\pi^*) \leq J(\pi) \ , \ \forall\ \pi^m \in \Pi^m, \ \ m \in \overline{M},$$

which implies the NM-tuple of inequalities

$$J(\pi^*) \leq J(\pi^{1^*};\ldots;\pi^{m-1^*};\gamma_1^{m^*},\ldots,\gamma_{n-1}^{m^*},\gamma_n^m,\gamma_{n+1}^{m^*},\ldots,\gamma_N^{m^*};\pi^{m+1^*};\ldots;\pi^{M^*}),$$

$$\forall \ \ \gamma_n^m \in \Gamma_n^m; \ n \in \overline{N}, \ m \in \overline{M}.$$

This set of inequalities (also known as person-by-person optimality, if we view each u_n^m to be controlled by a different agent) therefore provides a necessary condition for π^* to be a team-optimal solution. Note that here, all sub-strategies are held at their optimal values and the resulting cost functional is minimized over possible strategies $\gamma_n^m \in \Gamma_n^m$; hence each minimization problem is basically of the form

$$E \ \{ \min_{u_n^m} \ \int L_n^m(\xi,u_n^m) dP(\xi|\eta_n^m)\} \tag{4.6}$$

where

$$L_n^m(\xi,u_n^m) \ \triangleq \ L(\xi,\pi^{1^*}(\eta^1);\ldots;\gamma_1^{m^*}(\eta_1^m),\ldots u_n^m,\ldots,\gamma_N^{m^*};\ldots,\pi^{M^*}(\eta^M))$$

and $P_n^m(\xi|\eta_n^m)$ is the conditional probability distribution of the primitive random variables ξ given the information vector η_n^m. This conditional probability distribution is also known as the <u>sufficient statistics</u> for DM m at stage n. E{.} denotes the expected value over the statistics of η_n^m, after $u_n^m = \gamma_n^m(\eta_n^m)$ is determined. The reason why L_n^m can be determined explicitly

as a function of (ξ, u_n^m) is because the information structure is causal, and hence elimination of other variables by iterative substitution is possible.

Whenever η_n^m is partially nested and the measurements that appear in η_n^m satisfy the separability condition (4.5), the sufficient statistics have a simpler form which is basically static in nature. To see this, firstly construct (in view of Proposition 4.1) the largest partially nested information structure (say, $\bar{\eta}_n^m$) that is equivalent to η_n^m. This new information structure $\bar{\eta}_n^m$ clearly has the property that whenever $\eta_k^i \subseteq \bar{\eta}_n^m$ for any $k \leq n$, $i \in \bar{M}$, we have $u_k^i \in \bar{\eta}_n^m$. Because of separability of (4.5) and the partially nested property of $\bar{\eta}_n^m$, we have the further (sigma-field)equivalence

$$\bar{\eta}_n^m \equiv \tilde{\eta}_n^m$$

where $\tilde{\eta}_n^m$ is obtained from $\bar{\eta}_n^m$ by replacing all y_k^i with

$$\tilde{y}_k^i = \tilde{H}_k^i(x_1;\ \theta_n, \ldots, \theta_1;\ \theta_k^i).$$

Therefore,

$$P(\xi | \eta_n^m) = P(\xi | \bar{\eta}_n^m) = P(\xi | \tilde{\eta}_n^m).$$

But, since $\tilde{\eta}_n^m$ is also partially nested, the presence of the control values in $\tilde{\eta}_n^m$ does not provide any additional information, and we may as well consider the smaller set

$$\hat{\eta}_n^m = \tilde{\eta}_n^m \cap \{y_n^1, \ldots, y_n^M;\ y_{n-1}^1, \ldots, y_{n-1}^M;\ y_1^1, \ldots, y_1^M\}$$

which is totally static. Hence,

$$P(\xi|\tilde{\eta}_n^m) = P(\xi|\hat{\eta}_n^m),$$

which implies that there exists an equivalent static sufficient statistics for DM m at stage n. This leads to the following important conclusion.

Proposition 4.2.

(i) Any stochastic team problem with dynamic partially nested information structure $\{\eta_n^m, n \in \overline{N}, m \in \overline{M}\}$, whose measurement equations also satisfy the property (4.5), is equivalent to one with a static information structure $\{\hat{\eta}_n^m; n \in \overline{N}, m \in \overline{M}\}$ as constructed above, in the sense that the optimal solution of one can be obtained from the optimal solution of the other.

(ii) If $\{\hat{u}_n^{m*} = \gamma_n^{m*}(\hat{\eta}_n^m); n \in \overline{N}, m \in \overline{M}\}$ denotes the optimal team solution under the equivalent static information structure, the solution of the original team problem can be expressed as $\{u_n^{m*} = \gamma_n^{m*}(\eta_n^{m*}); n \in \overline{N}, m \in \overline{M}\}$ where η_n^{m*} is obtained from $\hat{\eta}_n^m$ by replacing $\tilde{y}_k^i$ with

$$y_k^i - \tilde{G}_k^i(u^1,\ldots,u^M)$$

and by appropriately replacing some of the controls with their optimum values, in a way compatible with the underlying information structure. [If the original information structure $_n^m$ is the largest partially nested information structure that is equivalent to itself, this latter phase is not required]. □

Remark. The separability condition (4.5) of the Proposition can be relaxed to some extent. The real requirement here is that the conditional probability

$P(\xi|\,{}^{m}_{n})$ should be independent of the control laws, so that there can be found a static information structure $\hat{\ }^{m}_{n}$ with the property $P(\xi|\,{}^{m}_{n}) = P(\xi|\,\hat{\ }^{m}_{n})$. A more relaxed condition [than (4.5)] that achieves this is given in Ho and Chu (1973). □

The result of Proposition 4.2 is very useful in stochastic team problems, because derivation of the optimal team solution under static information is in general much simpler than the derivation under dynamic information. In particular, for the special case when (i) the measurement equations are linear in the primitive random variables and the controls, (ii) the primitive random variables are jointly Gaussian distributed, (iii) the cost functional L is quadratic in the control vectors and the primitive random variables, and (iv) L is further strictly convex in the control variables, the unique team optimal solution is affine in the available information and can readily be computed by solving the set of minimization problems (4.6) [see Radner (1962), Ho and Chu (1972)]. Therefore, every linear-quadratic-Gaussian stochastic team problem with strictly convex cost functional and partially nested information structure admits a team-optimal solution that is affine in the available information —a result which directly follows from Radner's above mentioned result in view of Proposition 4.2. Furthermore, team-optimal control laws can be obtained recursively when the partially nested information pattern is one-step delay information sharing [Kurtaran (1975), Sandell and Athans (1974), Yoshikawa (1975)] or one-step-delay observation sharing [Başar (1978a)]. The solution is unique in the latter case and nonunique in the former case --the nonuniqueness arises because the one-step delay information sharing pattern includes redundant information which gives rise to several

different "representations" [see, Başar (1978a)].

If the underlying information structure in a stochastic team problem is nonclassical, derivation of the optimal team solution meets with formidable difficulties. Even in the simplest type of a linear-quadratic-Gaussian team problem with a two-step delay information sharing pattern (i.e. a nonclassical information pattern) the optimal solution is nonlinear and cannot be obtained analytically; moreover even a numerical derivation is a challenging task because such problems admit several person-by-person optimal solutions and local optima [see, Witsenhausen (1968)]. There are also no simple sufficient statistics for such problems with nonclassical information patterns [see,Yoshikawa and Kobayashi (1978), and Varaiya and Walrand (1978)]. These difficulties are due to the fact that each control has in general a "triple" role in stochastic team problems [Ho (1980)]: (i) the deterministic control effort of reducing the error, (ii) to improve the future knowledge of uncertainty, (iii) to "signal" the agents acting in the future some useful information which they will not necessarily acquire [in the case of classical or quasi-classical information patterns, this third role is absent]; and these three roles are in general in conflict with each other. Only if these roles are isolated, the stochastic team problems with nonclassical information patterns tend to be comparatively tractable [see, Witsenhausen (1975), and Ho, Kastner and Wong (1978)] --but this is indeed a very special class of problems and the more general nonclassical stochastic team problems await innovative ideas, techniques and approaches.

4.3 Nash Equilibria

Derivation of Nash equilibria for stochastic systems controlled by

several agents with different objective functionals is, in general, an extremely challenging problem when the information pattern is nonclassical --the reasons being similar to those we have discussed above at some length in the context of team problems. Therefore, we will confine our discussion in the sequel to deterministic, and stochastic classical and quasi-classical information patterns.

We have seen in §3.3 that in the case of deterministic systems with deterministic dynamic information patterns, there exists, in general, a multitude of Nash equilibria --the reason being that in such problems (i) every control law has several different "representations" and (ii) every Nash equilibrium obtained under an information structure that is inferior to the original deterministic information structure constitutes a Nash solution also under the original information structure. We call such equilibria "informationally nonunique" Nash solutions. For stochastic systems of the type (2.1), however, informationally nonunique Nash equilibria cannot occur, even under deterministic dynamic state information, provided that (roughly speaking) the noise vector θ_n "influences" all points in the state space X, and for every $n \in \overline{N}$ [Başar (1976, 1979a)]. A more precise statement can be given for the case when θ_n has an additive effect, that is when (2.1) is written as

$$x_{n+1} = f_n(x_n, u_n^1, \ldots, u_n^M) + \theta_n. \tag{5.7}$$

The requirement here is that the probability measure $\mathcal{P}_n$ associated with θ_n should assign positive probability to every open subset of X [assuming that an appropriate topology is defined on X] [Van Damme(1980)]. Such a stochastic formulation ensures existence of a unique representation for every strategy and

hence eliminates the possibility of having informationally nonunique Nash equilibria under dynamic state information (such as the closed-loop perfect state information). The only nonuniqueness (if any) will be due to the structures of the cost functions and the state equation.

Consider now the case when the system equation is given by (4.7)[with the probability measure of θ_n having the property discussed above], the underlying information structure is closed-loop perfect state, and the cost functional of DM m ($m \in \overline{M}$) is given by

$$J^m = E\{ \sum_{n=1}^{N} g_n^m(x_{n+1}, u_n^1, .., u_n^M, x_n) \mid u_n^m = \gamma_n^m(\eta_n^m),\ n \in \overline{N},\ m \in \overline{M}\} .$$

For such problems the Nash equilibrium solution can be computed recursively, by following a dynamic programming type argument and by solving at each stage a static Nash problem. Assuming that each $f_n(.)$ and g_n^m ($n \in \overline{N}$, $m \in \overline{M}$) is continuously differentiable in its arguments, and $\{\theta_n,\ n \in \overline{N}\}$ is an independent sequence, the recursive relation that yields the Nash solution $\{ u_n^{m^*} = \gamma_n^{m^*}(x_n);\ n \in \overline{N},\ m \in \overline{M}\}$ reads [c.f. Başar (1979a)]:

$$\int_X \{\nabla_m f_n(x_n, u_n^{1^*}, .., u_n^{M^*}) \nabla_o G_n^m(x_{n+1}^*, u_n^{1^*}, .., u_n^{M^*}, x_n) + \nabla_m G_n^m(x_{n+1}^*, u_n^{1^*}, .., u_n^{M^*}, x_n) d\theta_n = 0$$

$$G_n^m(x_{n+1}, u_n^1, .., u_n^M, x_n) = \int_X G_{n+1}^m(x_{n+2}^*, u_{n+1}^{1^*}, .., u_{n+1}^{M^*}, x_{n+1}) d\theta_{n+1}$$

$$+ g_n^m(x_{n+1}, u_n^1, .., u_n^M, x_n)$$

$$G_N^m \equiv 0, \quad m \in \overline{M}, \quad n \in \overline{N}$$

where

$$x^{*}_{n+1} \triangleq f_n(x_n, u_n^{1*}, \ldots, u_n^{M*}) + \theta_n$$

and $\mathcal{P}_n$ denotes the probability measure of θ_n. It should be noted that "informational nonuniqueness" is absent here, mainly because of our assumption on $\mathcal{P}_n$ $(n \in \overline{N})$, and it is for this reason that every solution set will be a function of only the current value of the state. When the state equation is linear and each cost functional is quadratic, a unique solution can be obtained under some invertability conditions on system matrices, and the Nash control laws are affine functions of the current values of the state (depending only on the mean value of θ_n) [Başar (1979a)].

When the underlying information structure is quasi-classical, derivation of the Nash equilibrium solution is a more subtle issue. Firstly, Propositions 4.1 and 4.2 do not have any counterparts here, which totally removes the possibility of simplifying the information structure (such as reconsidering the original problem under an "equivalent" static information). Secondly, if the underlying information pattern is the one-step delay information sharing pattern, there exists, in general, a plethora of "informationally nonunique" Nash equilibria, because that particular information pattern incorporates redundancy in dynamic information [each agent having access to past measurements as well as to past control values of the other agents].[See Başar (1978a) for a class of such informationally nonunique Nash equilibria.] In order to avoid informationally nonunique equilibria, we have to restrict our attention to those quasi-classical information patterns which are free of any redundancy in dynamic information --such as the one-step-delay observation sharing pattern.

The derivation of Nash equilibria under the one-step-delay observation sharing pattern is not a totally intractable problem, and, depending on the structures of the cost functions, measurement equations and state equation, one can utilize a partially recursive procedure (of the dynamic programming type) that would yield the optimal solution. This procedure (whenever it works) involves, at each stage, the solution of static stochastic Nash problems and satisfaction of some consistency conditions; however, as a caveat for the reader we should mention that such a derivation is not routine and it involves several pitfalls, mainly due to the fact that the conditional distribution of the state at each stage (given the past and present acquired information) depends in general on the past control laws [hence, the derivations at each stage cannot totally be isolated, as in the case of stochastic team problems discussed in §4.2].

Let us now outline this procedure in some general terms, by pointing out the difficulties as they arise. Suppose that the Nash equilibrium solution has been determined up to the last stage, and we are faced with the "static" last stage Nash problem which has the cost function

$$J_N^m = E\{g_N^m[f_N(x_N,u_N^1,\ldots,u_N^M) + \theta_N, u_N^1,\ldots,u_N^M, x_N] \mid u_N^i = \gamma_N^i(\eta_N^i),\ i \in \overline{M}\}$$

for DM m, where the probability distribution of x_N depends on the past controls through the state equation (4.7). Denote the Nash solution of this problem by

$$\gamma_N^m(\eta_N^m) = \varphi_N^m(y_N^m; y_{N-1}; \ldots; y_1), \quad m \in \overline{M}. \tag{4.8}$$

[Derivation of this solution will in general be quite difficult; however, the difficulty is not a conceptual one but rather a computational one. We will discuss this point further, in the sequel, for the special case of linear-quadratic problems.] Here, φ_N^m will depend on the conditional probability distribution of x_N, and thereby on the past control laws. Now, if the structural dependence of φ_N^m on y_N^m depends explicitly on the past control laws, the procedure cannot be carried over to the next stage, since y_N^m also depends on $\{u_{N-1}^i;\ i \in \overline{M}\}$ and therefore the general structure of the Nash problem at stage N-1 will depend (implicitly) on the solution that is being sought. This difficulty can, however, be avoided, if (4.8) happens to be separable in y_N^m, i.e.

$$\varphi_N^m(y_N^m, y_{N-1}, \ldots, y_1) = \hat{\varphi}_N^m(y_N^m) + \tilde{\varphi}_N^m(y_{N-1}, \ldots, y_1) \tag{4.9}$$

with the further property that $\hat{\varphi}_N^m$ is functionally independent of the past control laws. In such a case, the dependance of the Nash equilibrium strategies at stage N on the controls at stage N-1 [i.e. $\{u_{N-1}^m,\ m \in \overline{M}\}$] are completely determined by the functions $\{\hat{\varphi}_N^m,\ m \in \overline{M}\}$], and therefore we can proceed to the next stage (N-1) for determination of the Nash control laws $\{\gamma_{N-1}^{m*},\ m \in \overline{M}\}$ by substituting

$$u_N^m = \bar{u}_N^m = \hat{\varphi}_N^m(y_N^m) + k_N^m(y_{N-1}, \ldots, y_1), \quad m \in \overline{M}, \tag{4.10}$$

in the state equation and the cost functionals, where k_N^m is any measurable

function of its arguments. The "static" Nash problem (of interest) at stage N-1 will then involve the cost functionals

$$J^m_{N-1} = E\{g^m_{N-1}[f_{N-1}(x_{N-1},u^1_{N-1},\ldots,u^M_{N-1}) + \theta_{N-1},u^1_{N-1},\ldots,u^M_{N-1},x_{N-1}]$$

$$+ g^m_N[f_N(x_N,\bar{u}^1_N,\ldots,\bar{u}^M_N) + \theta_N,\bar{u}^1_N,\ldots\bar{u}^M_N,x_N] \mid u^i_{N-1} = \gamma^i_{N-1}(\eta^i_{N-1}),$$

$$i \in \bar{M}\}, \quad m \in \bar{M},$$

where $\bar{u}^m_N$ is given by (4.10), x_N is related to the controls at stage N-1 through

$$x_N = f_{N-1}(x_{N-1},u^1_{N-1},\ldots,u^M_{N-1}) + \theta_{N-1},$$

and y^m_N is related to the past controls through

$$y^m_N = h^m_N(x_N,\theta^m_N).$$

Now suppose that, for a fixed set of sub-strategies at stages N-2,N-3,..,1, the solution of this Nash problem exists and is given by

$$\gamma^m_{N-1}(\eta^m_{N-1}) = \hat{\varphi}^m_{N-1}(y^m_{N-1}) + \bar{\varphi}^m_{N-1}(y_{N-2},\ldots,y_1), \; m \in \bar{M}, \tag{4.11}$$

where $\hat{\varphi}^m_{N-1}$ is functionally independent of the past control laws, but it may depend on $\{k^i_N,\ i \in \bar{M}\}$ which in turn depends on the value of γ^m_{N-1} at equilibrium through the second term in (4.9). Invoking the consistency condition, and re-solving for γ^m_{N-1} from (4.11), we obtain the structural form

$$\gamma^m_{N-1}(\eta^m_{N-1}) = \hat{\hat{\varphi}}^m_{N-1}(y^m_{N-1}) + \bar{\bar{\varphi}}^m_{N-1}(y_{N-2},\ldots,y_1), \; m \in \bar{M},$$

where $\hat{\phi}^m_{N-1}$ does not depend on either the past controls or $\{k^i_N, i \in \bar{M}\}$. Hence, we can now let

$$u^m_{N-1} = \bar{u}^m_{N-1} = \hat{\phi}^m_{N-1}(y^m_{N-1}) + k^m_{N-1}(y_{N-2},\ldots,y_1), \; m \in \bar{M},$$

where k^m_{N-1} is any measurable function of its arguments, and repeat the deeds of stage N-1 at stage N-2. Then, the solution can be obtained inductively by invoking the consistency requirement at every stage, under the assumptions that at every stage a Nash equilibrium solution to the related static problem exists, and it satisfies a separability condition of the type (4.9) or (4.11).

The above outlined procedure has been implemented in Başar (1978b) for the class of linear-quadratic-Gaussian (LQG) systems under the one-step-delay observation sharing pattern, and existence of a unique Nash solution, linear in the available information, has been verified under some sufficiency conditions that involve the system parameters. The "static" stochastic Nash problem to be solved at each stage is of the linear-quadratic type, whose solution is discussed in Başar (1975) and Başar (1978a), which may be considered as an extension of Radner's result [Radner (1962)] referred to in §4.2 to problems with different objectives for different agents. We should mention that the solution of the general LQG problem given in Başar (1978b) is highly complicated in terms of the equations which yield the coefficient matrices of the linear control laws, and it does not satisfy any separation property (as opposed to the solution of the LQG team problem under the same information pattern).

When the underlying information structure is nonclassical, derivation of the Nash equilibrium solution is <u>in general</u> not tractable, since even the

special case of nonclassical team problems involve formidable difficulties, as discussed earlier in §4.2. However, there exists a subclass of problems with totally conflicting goals, whose Nash equilibrium solutions (rather called saddle-point solutions in this context) can be obtained explicitly (and analytically) even under nonclassical information patterns, mainly because in such problems controls of the agents do not have "triple" role (i.e. the signaling aspect is absent). For example, Witsenhausen's counter example [Witsenhausen (1968)], when cast in such a framework, admits unique Nash (saddle-point) equilibrium that is linear in the available nonclassical information [see, Başar and Mintz (1972)]. For more discussion on such solvable stochastic problems with nonclassical information patterns, see Başar and Mintz (1971, 1973).

4.4. Hierarchical Decision Structure

In this subsection, we discuss the problem of optimal control and coordination of stochastic systems under hierarchical decision structure, by employing the Stackelberg solution concept introduced in §2.2 and elaborated on in §3.4 for deterministic systems. Let us first direct our attention to the case of two agents with different goals, and with DM 1 (called the leader) being in a position to enforce his strategy on DM 2 (known as the follower).

Information structure again plays a crucial role in such problems, and solvability of a specific problem depends to a great extent on the nature of the underlying information pattern. We should mention, at the outset, that stochastic decision problems in which the leader has access to static or dynamic redundant information (such as the one-step delay information sharing pattern) are much more tractable as compared with those in which the leader

has dynamic (non-redundant) information (such as the one-step delay observation sharing pattern) --this latter class of problems is in fact extremely challenging and as to date no general method exists that would aid in their solution.

Static information

When the leader has access to static information [more precisely, if the leader's information does not depend on the controls of the follower], the stochastic Stackelberg problem is tractable because the rational response set of the follower does not structurally depend on the strategy of the leader. Such problems are then essentially equivalent to one-stage stochastic Stackelberg problems[†] which we now discuss. In terms of the standard notation, let

$$J^m = E\,\{g^m(u^1,u^2,\xi) \mid u^i = \pi^i(\eta^i),\ i = 1,2\},\ m = 1,2,$$

where

$$\eta^i = \{y^i\}\ ,\quad y^i = h^i(\xi),\quad i = 1,2,$$

and ξ denotes a collection of primitive random variables with known probability distributions. Let $\pi^1 \in \Pi^1$ be fixed, where Π^1 is appropriately defined. Then, the follower is faced with the stochastic minimization problem

$$\begin{aligned} &\min_{\pi^2 \in \Pi^2} E\,\{g^2[\pi^1(h^1(\xi)),\ \pi^2(y^2),\xi]\} \\ &\quad = \min_{u^2} E\,\{g^2[\pi^1(h^1(\xi)),\ u^2,\xi] \mid y^2\}, \end{aligned} \tag{4.12}$$

[†]When the follower has access to dynamic information, there is no loss of generality in replacing it with an appropriate static information.

whose structure does not depend on the choice of π^1 since π^1 does not carry u^2 in its argument. If g^2 is strictly convex in u^2, this minimization problem admits a unique solution [regardless of the choice of π^1] which we denote by $T: \Pi^1 \to \Pi^2$, so that $\pi^2 = T\pi^1$ uniquely solves (4.12). The Stackelberg strategy π^{1*} is then any solution of the stochastic minimization problem

$$\begin{aligned} &\min_{\pi^1 \in \Pi^1} E\,\{g^1[\pi^1(y^1),\ T\pi^1(y^1), \xi]\} \\ &\quad = \min_{u^1} E\,\{g^1[u^1, T u^1, \xi] \,|\, y^1\}. \end{aligned} \tag{4.13}$$

The two optimization problems (4.12) and (4.13) can be solved (at least numerically) without any major difficulty of conceptual or methodological nature, and in a few cases the solution can be obtained analytically. One such specific case is the class of linear-quadratic-Gaussian systems [g^1 and g^2 quadratic, h^1 and h^2 linear, and ξ Gaussian], for which the Stackelberg solution is affine. More precisely, we have from Başar (1979a, 1980a).

<u>Proposition 4.3</u>.

Let $\xi = (x, \theta^1, \theta^2)$ be Gaussian distributed with mean zero and covariance $\mathrm{diag}(\Sigma, \Lambda^1, \Lambda^2)$. Further let

$$g^m(x,u^1,u^2) = \frac{1}{2}u^{m'}D_{mm}u^m + u^{m'}D_{mi}u^i + \frac{1}{2}u^{i'}D_{m3}u^i + u^{m'}C_{mm}x$$

$$+ u^{i'}C_{mi}x; \quad m,i = 1,2; \quad i \neq m, \quad D_{mm} > 0,$$

and $\qquad h^m(x,\theta^m) = H^m x + \theta^m, \qquad m = 1,2.$

Then, the stochastic Stackelberg problem with static information

$\eta^m = \{y^m\}$, m = 1,2, admits the <u>unique</u> solution

$$\gamma^{1^*}(y^1) = Ay^1; \gamma^{2^*}(y^2) = -D_{22}^{-1}[C_{22}\Sigma_2 y^2 + D_{21}E[\gamma^{1^*}(y^1) | y^2]]$$

where A is the <u>unique</u> solution of the Lyapunov-type equation

$$D_{11}A + [D_{21}'D_{22}^{-1}D_{13}D_{22}^{-1}D_{21} - D_{12}D_{22}^{-1}D_{21} - D_{21}'D_{22}^{-1}D_{12}']\, A\,\Sigma_2 H^2 \Sigma_1$$

$$= [(D_{12} - D_{21}'D_{22}^{-1}D_{13})D_{22}^{-1}C_{22} + D_{21}'D_{22}^{-1}C_{12}]\Sigma_2 H^2 \Sigma_1 - C_{11}\Sigma_1$$

and

$$\Sigma_m \triangleq \Sigma H^{m'}(H^m \Sigma H^{m'} + \Lambda^m)^{-1}, \quad m = 1,2,$$

<u>provided that</u> the condition

$$0 < I + D_{11}^{-1/2}[D_{21}'D_{22}^{-1}D_{13}D_{22}^{-1}D_{21} - D_{12}D_{22}^{-1}D_{21} - D_{21}'D_{22}^{-1}D_{12}']\, D_{11}^{-1/2} < 2I$$

holds. □

<u>Remark</u>. The preceding result may be considered as an extension of Radner's result on LQG teams, cited in §4.2, to problems with different objective functionals for the agents and with a hierarchical decision structure. Even though this specific result pertains to the two-agent case, its extensions to the multi-agent case with more than two levels of hierarchy (in decision making) can be envisioned -- such problems (when cast in the LQG framework) also admit unique affine solutions, but the verification and the derivation are much more complicated than in the two-agent case [Başar (1981c)]. Yet another extension (and application) of Proposition 4.3 would be to dynamic

decision problems under the feedback (stagewise) Stackelberg solution concept, in which case the leader enforces his strategy on the follower only stagewise. For LQG dynamic systems and under the one-step-delay observation sharing pattern, it can be shown by repeated application of Proposition 4.3, together with a dynamic programming type argument, that the feedback Stackelberg solution is affine in the information available to the two agents [see, Başar (1979)]. □

Dynamic redundant information

We have earlier discussed in §4.3 that presence of redundancy in the dynamic information structure gives rise to ill-posed problems in the case of Nash equilibria, because it leads to a plethora of informationally nonunique solutions. For problems with hierarchical decision structure, however, the situation is quite the opposite. This time, presence of redundancy in the dynamic information actually helps to simplify the derivation of the Stackelberg solution, because the extra freedom allotted to the leader through the redundancy enables him to provide incentives or implement threats for the follower, so as to force him to the most favorable solution [from the leader's point of view]. We have already elucidated this property of redundant dynamic information in §3.4 for deterministic systems, and in the following we discuss it for stochastic systems within the context of a specific model.

Consider the general two-agent decision problem treated earlier in this subsection, but under the amended information structure

$$\eta^1 = \{y^1, y^2, u^2\}, \qquad \eta^2 = \{y^2\},$$

that is, the leader has also access to the measurement and control value of

the follower. [Of course, this makes sense only if the follower acts before the leader does, which we assume to be valid in this case]. Now let

$$\min_{\pi^1 \in \Pi^1,\, \pi^2 \in \Pi^2} E\,\{g^1[\xi, \pi^1(y^1,y^2),\ \pi^2(y^2)]\}$$

exist and be determined uniquely by

$$u^{1^t} = \pi^{1^t}(y^1,y^2), \qquad u^{2^t} = \pi^{2^t}(y^2).$$

Let $E\,\{g^2(\xi, u^{1^t}, u^{2^t})\} = g^{2^t}$, and let there exist a $\bar{\pi}^1 \in \Pi^1$ such that

$$\min_{\pi^2 \in \Pi^2} E\,\{g^2[\xi, \bar{\pi}^1(y^1,y^2,u^2), u^2] \mid u^2 = \pi^2(y^2)\} > g^{2^t}.$$

Then, by announcing the strategy

$$\pi^{1^*}(\eta^1) = \begin{cases} u^{1^t} & \text{if } u^2 = \pi^{2^t}(y^2) \\ \bar{\pi}^1 & \text{otherwise} \end{cases},$$

the leader can force the follower to adopt the strategy π^{2^t}, and thereby incur an overall favorable cost value. We can therefore declare π^{1^*} as a Stackelberg strategy for the leader and consider the problem solved. However, for several reasons, one may wish to replace the essential threat π^{1^*} with a "softer" incentive scheme which penalizes the follower proportionately to his deviation from the desired solution. Such incentive schemes (which are basically different representations of π^{1^*}) do exist, and for several discussions and derivations, as well as on extensions of this approach, we refer to Başar

(1980a), and also to Ho, Luh and Mulidharan (1981). Extensions to the case of multi-levels of hierarchy are discussed in Başar (1981a).

Non-redundant dynamic information

For the procedure outlined above to work, the information structure of the leader should be such that if, at stage $n \in \bar{N}$, η_n^1 depends on u_k^2 for some $k < n$ [either directly or through the measurement equation], then he should know both the value of u_k^2 and the information η_k^2 on which it is based. With such an information structure, which incorporates redundancy, the leader can, in general, enforce the solution that is most favorable to him. If the information structure is dynamic, but does not incorporate any redundancy, the Stackelberg solution is extremely difficult to obtain, unless one parameterizes the desired solution and converts the original dynamic optimization problem to a static one (over those parameter values). Such an approach, of course, leads in general to suboptimal Stackelberg solutions. Even for linear-quadratic stochastic systems with perfect state measurements, there is no known method to obtain the closed-loop Stackelberg solution, and the linear suboptimal solution can only be obtained numerically, with the coefficient matrices depending on the statistical parameters of the additive system noise [see, Başar (1979a)].

The following table [Table 1] now recapitulates, in a nutshell, the known results and the yet-unsolved problems in the control and coordination of stochastic systems with multiple decision makers and under different types of information, together with related references. We have classified the problems in four categories.

(1) Completely solved ones --remaining details are of minor nature.

(2) Not completely solved. Any new result on this class of problems

will constitute a contribution to the field, but a totally innovative approach is not required.

(3) Some "positive" and "negative" partial results on special cases exist; but this general class of problems is extremely challenging, and innovative approaches have to be introduced in order to solve a sufficiently general class of such problems.

(4) These problems are ill-posed, mainly because they lead to a plethora of solutions which cannot be strictly ordered.

The references quoted in the Table are not meant to be exhaustive; we have chosen to list the most recent or the most representative ones in each subcategory.

Information Structure		Team	Saddle-point	Nash	Stackelberg
Closed-loop perfect state	LQG	(1) Bertsekas (1976)	(1) Başar and Olsder (1982)	(1) Başar and Olsder (1982)	(3) Cruz (1978)
	other	(2) "	(2) "	(2) "	
Stochastic Centralized	LQG	(1) Bertsekas (1976)	(1) Başar (1981b)	(2) Başar (1978b)	(3) Cruz (1978)
	other	(2) "	(2) "	(2)	
1-step delay observation sharing	LQG	(1) Başar (1978a)	(2) Başar (1978b)	(1) Başar (1978a,b)	(3) Başar (1979b)
	other	(2)		(2)	
1-step delay information sharing	LQG	(1) Sandell and Athans (1974), Ho (1980	(2) Başar (1978b)	(4)	(2) Başar (1980a, 1981c)
	other	(2)			
Partially nested non-redundant	LQG	(1) Ho and Chu (1972, 1973), Ho (1980)	(2)	(2)	(3)
	other	(2)			
Partially nested redundant	LQG	(1) Ho and Chu (1972, 1973), Ho (1980)	(2)	(4)	(2) Başar (1981c)
	other	(2)			
Strictly Nonclassical		(3) Witsenhausen (1968)	(2) Başar and Mintz (1971, 1972, 1973)	(3)	(3)

TABLE 1: A display of the current "state of Knowledge" on the control and optimization of discrete-time stochastic systems.

5. APPLICATIONS

In this section we consider a few simplified situations where the concepts of multiperson decision-making are meaningful. The examples are intended to suggest potential areas where the concepts may be used as guides in decision-making.

5.1. Nash Equilibrium Model of an Arms Race

Richardson's model [Richardson (1960)] of arms race between two nations:

$$\dot{x}_1(t) = \sigma x_2(t) - \alpha x_1(t) + g \tag{5.1}$$

$$\dot{x}_2(t) = \rho x_1(t) - \gamma x_2(t) + h \tag{5.2}$$

has generated some interest in political science in further exploration of mathematical models in international relations. The arms levels (at time t) of two nations are represented by $x_1(t)$ and $x_2(t)$, σ and ρ are called defense coefficients, α and γ are called fatigue coefficients, and g and h are grievance coefficients. Discretizing time, the model may be represented in multistage form as

$$x_1(k+1) = a_{12}(k)x_2(k) + a_{11}(k)x_1(k) + b_1(k) \tag{5.3}$$

$$x_2(k+1) = a_{21}(k)x_1(k) + a_{22}(k)x_2(k) + b_2(k). \tag{5.4}$$

In an attempt to model how the coefficients $a_{ij}(k)$ might evolve and to attempt to explain how the nations' decision processes might lead to the model in (5.1) and (5.2), Simaan and Cruz [Simaan and Cruz (1975a)] proposed a Nash equilibrium model for the following multiperson decision problem:

The fundamental model for the arms levels is given by the pair of equations:

$$x_1(k+1) = \beta_1 x_1(k) + Z_1(k) \tag{5.5}$$

$$x_2(k+1) = \beta_2 x_2(k) + Z_2(k) \tag{5.6}$$

where $\beta_1 x_1(k)$ and $\beta_2 x_2(k)$ are the depreciated values of the arms stocks at stage k+1, and $Z_1(k)$ and $Z_2(k)$ are investments in arms. We seek strategies which are feedback Nash equilibrium strategies with respect to some objective functions. Thus $Z_1(k)$ and $Z_2(k)$ will be functions of the current arms levels $x_1(k)$ and $x_2(k)$. The objective functions are modeled to be

$$\begin{aligned} J_i(Z_1,Z_2) &= \frac{1}{2} Q_i(N+1)(x_i(N+1) - P_i(N+1)x_j(N+1) - V_i(N+1))^2 \\ &+ \frac{1}{2} \sum_{k=1}^{N} \{R_i(k)(Z_i(k)-W_i(k))^2 \\ &+ Q_i(k)(x_i(k)-P_i(k)x_j(k)-V_i(k))^2\}, \qquad i=1,2, \end{aligned} \tag{5.7}$$

where $R_1(k)$ and $R_2(k)$ are strictly positive real numbers and $Q_1(k)$, $Q_2(k)$, $P_1(k)$, and $P_2(k)$ are nonnegative real numbers for each k. Thus each nation wishes to narrow the gap between its armament level and an affine function of its opponent's armament level,while at the same time minimizing its armament expenditures.

Using dynamic programming the feedback Nash equilibrium solutions are found to be

$$Z_1(n) = A_{11}(n)x_1(n) + A_{12}(n)x_2(n) + B_1(n) \tag{5.8}$$

$$Z_2(n) = A_{21}(n)x_1(n) + A_{22}(n)x_2(n) + B_2(n) \tag{5.9}$$

where $A_{ij}(n)$ and $B_i(n)$ satisfy some recursive equations. When substituted in (5.5) and (5.6), the final feedback Nash equilibrium model is given by

$$x_1(k+1) = (\beta_1 + A_{11}(k))x_1(k) + A_{12}(k)x_2(k) + B_1(k) \tag{5.10}$$

$$x_2(k+1) = (\beta_2 + A_{22}(k))x_2(k) + A_{21}(k)x_1(k) + B_2(k). \tag{5.11}$$

Thus the coefficients in the discrete-time Richardson model of (5.3) and (5.4) may be related to the depreciation coefficients in (5.5) and (5.6), and to the coefficients of the objective functions in (5.7) associated with a multi-person decision problem. Thus the modeling problem is shifted to a choice of weighting coefficients in the objective functions of (5.7). For more details see Simaan and Cruz (1975a). An outline for obtaining the feedback Stackelberg solution for this arms race problem is given in Simaan and Cruz (1976).

5.2. Dynamic Duopoly with Production Constraints

In Simaan and Takayama (1978), a dynamic duopoly model with a linear demand of the form

$$\dot{p} = C - ap - b(x_1 + x_2) \tag{5.12}$$

where p is the commodity price and x_i is the output of firm i. The cost of production is

$$g_i(x_i) = \frac{1}{2}\alpha_i x_i^2, \qquad i = 1,2\ , \tag{5.13}$$

and the total profit for firm i over the horizon T is

$$\Pi_i(x_1, x_2) = \int_0^T \exp(-rt)[px_i - \frac{1}{2}\alpha_i x_i^2]dt \tag{5.14}$$

for i=1,2. The productions x_i are to be chosen as functions of the instantaneous price p(t) and it is assumed that the production capacity constraints are

$$0 \leq x_i[t,p(t)] \leq X_i, \qquad i=1,2. \tag{5.15}$$

Open-loop and feedback Nash equilibrium solutions are investigated in Simaan and Takayama (1978), where nine possibilities are explored, depending on whether firm i is not producing, producing at maximum capacity, acting as a monopolist, or playing as a true duopolist. For more details, see Simaan and Takayama (1978).

5.3. Electricity Pricing

Consider a simple model for electricity pricing, where the consumer chooses a level of consumption q to maximize his "consumer surplus" which is affected by the price of electricity. The electric utility chooses the revenue function r(q) to maximize its profit subject to capacity and subject to regulation. Such a problem was considered as a Stackelberg problem, with the utility as leader and the consumer as follower, by Ho, Luh, and Muralidharan (1981). Let the consumer surplus be modeled by

$$J_F = \frac{1}{2} S[\bar{q}^2 - (q-\bar{q})^2] - r(q) \tag{5.16}$$

where S and $\bar{q}$ are positive constants, r(q) is a monotonic increasing piece-wise linear function representing cost to the consumer (revenue to the utility). The profit of the utility is

$$J_L = r(q) - \frac{1}{2} cq^2, \tag{5.17}$$

the capacity constraint is

$$q \leq \hat{q}, \tag{5.18}$$

and the regulation constraint is

$$J_L \leq kq \tag{5.19}$$

where c, $\hat{q}$, and k are positive constants. The information structure is

$$\eta_F : \text{ no information}$$

$$\eta_L : q.$$

Ho, Luh, and Muralidharan (1981) determined that

$$r(q) = pq + F \tag{5.20}$$

is a Stackelberg strategy, where

$$p = S(\bar{q}-\hat{q}) \geq 0 \tag{5.21}$$

$$F = k\hat{q} + \frac{1}{2} c\hat{q}^2 - S\hat{q}(\bar{q}-\hat{q}). \tag{5.22}$$

The solution in (5.20), (5.21), and (5.22) has the property that J_L is maximized with respect to r and q. Furthermore, with r(q) given as in (5.20), the optimum value of q for the consumer is $\hat{q}$, the capacity of the utility. The resulting value of the utility profit, J_L, is $k\hat{q}$, which is the maximum allowed by regulation.

6. CONCLUDING REMARKS

In this chapter we discussed some key concepts and methods relevant to multi-person decision-making and optimization in dynamic systems. In large scale physical models, dynamic operations research models, and policy and planning models, it is important and crucial to explicitly model the roles of multiple decision makers if, indeed, there is more than one entity that makes choices. For certain purposes, such as in policy analysis, it may be adequate to recognize only one decision maker and subsume other decision-making aspects in general sectors. However, in the investigation of effects of significant policy changes, based on a model calibrated from data on previous policies, the predicted outcome may be misleading because when the policy is changed, the reactions of the subsumed decision makers may change so that the fixed model being used may not be satisfactory anymore. It would be preferable to explicitly model the presence of the other decision makers.

For situations where cooperation among decision-makers is desirable, the concept of Pareto optimality is appropriate. However, in non-cooperative situations the Nash equilibrium concept is more natural. Hierarchies in decision-making lead to the concept of Stachelberg or leader-follower strategies. These concepts are described in this chapter for both deterministic and stochastic systems.

A critical consideration in multi-person optimization problems is the information structure. In contrast to single person decision making which necessarily involves centralized information, the multi-person decision-making problem may involve decentralized information. Furthermore, the

assumption of memory in the measurement, even in the deterministic case, generally leads to a solution different from that with no-memory in the multi-person case. In contrast, memory in the measurement has no effect on the optimal solution for single person optimization problems.

For simplicity in exposition, only the class of discrete-time dynamic systems is treated. The concepts discussed in the chapter are also applicable to continuous-time dynamic systems.

ACKNOWLEDGMENT: Some of the research reported herein, and the writing of this chapter were supported in part by the Joint Services Electronics Program under Contract N00014-79-C-0424; in part by the National Science Foundation under Grant ECS-79-19396; and in part by the U.S. Air Force under Grant AFOSR-78-3633.

BIBLIOGRAPHY

[1] Bagchi, A. and T. Başar (1980), "Team decision theory for linear continuous-time systems", IEEE Transactions on Automatic Control, vol AC-25, no. 6, pp. 1154-1160.

[2] Bagchi, A. and T. Başar (1981), "Stackelberg strategies in linear-quadratic stochastic differential games", Journal of Optimization Theory and Applications, vol. 35, no. 3, pp. 343-364.

[3] Başar, T. (1973), "On the relative leadership property of Stackelberg strategies", Journal of Optimization Theory and Applications, vol. 11, no. 6, pp. 655-661.

[4] Başar, T. (1974), "A counter example in linear-quadratic games: Existence of nonlinear Nash solutions", Journal of Optimization Theory and Applications, vol. 14, no. 4, pp. 425-430.

[5] Başar, T. (1975a), "Equilibrium solutions in two-person quadratic decision problems with static information structures", IEEE Transactions on Automatic Control, vol. AC-20, pp. 320-328.

[6] Başar, T. (1975b), "Nash strategies for M-person differential games with mixed information structures", Automatica, vol. 11, no. 5, pp. 547-551.

[7] Başar, T. (1976), "On the uniqueness of the Nash solution in linear-quadratic differential games", International Journal of Game Theory, vol. 5, no. 2/3, pp. 65-90.

[8] Başar, T. (1977a), "Two general properties of the saddle-point solutions of dynamic games", IEEE Trans. on Automatic Control, vol. AC-22, no. 1, pp. 124-126.

[9] Başar, T. (1977b), "Informationally nonunique equilibrium solutions in differential games", SIAM Journal on Control and Optimization, vol. 15, no. 4, pp.636-660.

[10] Başar, T. (1977c), "Existence of unique Nash equilibrium solutions in nonzero-sum stochastic differential games", in: E. O. Roxin, P. T. Liu, R.L. Sternberg, eds., Differential Games and Control Theory II, Marcel Dekker, New York, pp. 201-228.

[11] Başar, T. (1978a), "Decentralized multicriteria optimization of linear stochastic systems," IEEE Transactions on Automatic Control, vol. AC-23, no. 3, pp. 233-243.

[12] Başar, T. (1978b), "Two-criteria LQG decision problems with one-step delay observation sharing pattern," Information and Control, vol. 38, no. 1, pp. 21-50.

[13] Başar, T. (1979a), "Information structures and equilibria in dynamic games," in: M. Aoki, A. Marzollo, eds., New Trends in Dynamic System Theory and Economics, Academic Press, New York, pp. 3-55.

[14] Başar T. (1979b), "Stochastic stagewise Stackelberg strategies for linear-quadratic systems,", in: M. Kohlmann, W. Vogel, eds., Stochastic Control Theory and Stochastic Differential Systems, Lecture Notes in Control and Information Sciences, vol. 16, Springer Verlag, pp. 264-276.

[15] Başar, T. (1979c), "Stochastic linear quadratic nonzero-sum differential games with delayed information patterns," in: A. Niemi, ed., Proceedings of the 7th IFAC World Congress, vol. 2, Pergamon Press, New York.

[16] Başar, T. (1980a), "Hierarchical decision making under uncertainty," in P.T. Liu ed., Dynamic Optimization and Mathematical Economics, Plenum Press, N.Y., pp. 205-221.

[17] Başar, T. (1980b), "Optimum coordination of linear interconnected systems," Large Scale Systems, vol. 1, no. 1, pp. 17-28.

[18] Başar, T. (1980c), "Memory strategies and a general theory for Stackelberg games with partial dynamic information," in: Proceedings of the Fourth International Conference on Analysis and Optimization of Systems, Versailles, France.

[19] Başar, T. (1980d), "On the existence and uniqueness of closed-loop sampled-data Nash controls in linear-quadratic stochastic differential games," in K. Iracki, K. Malonowski, S. Walukiewicz, eds., Optimization Techniques, part I, Lecture Notes in Control and Information Sciences, vol. 22, Springer Verlag, Berlin, pp. 193-203.

[20] Başar, T. (1980e), "Equilibrium strategies in dynamic games with multi levels of hierarchy," in: Proceedings of the Second IFAC Symposium on Large Scale Systems: Theory and Applications, Toulouse, France.

[21] Başar, T. (1981a), "Stochastic multicriteria decision problems with multi-levels of hierarchy," IEEE Transactions on Automatic Control, vol. AC-26, no. 2, pp. 549-553.

[22] Başar, T. (1981b), "On the saddle-point solution of a class of stochastic differential games," Journal of Optimization Theory and Applications, vol. 33, no. 4, pp. 539-556.

[23] Başar, T. (1981c), "Hierarchical equilibrium solutions in stochastic decision problems defined on general Hilbert spaces," submitted for publication.

[24] Başar, T. (1981d), "A new method for the Stackelberg solution of differential games with sampled-data state information," Proc. 8th IFAC Congress, Kyoto, Japan.

[25] Başar, T. (1981e), "Equilibrium strategies in dynamic games with multi-levels of hierarchy," Automatica, vol. 17, no. 5.

[26] Başar, T. (1981f), "Performance Bounds for hierarchical systems under partial dynamic information," Journal of Optimization Theory and Applications, to appear.

[27] Başar, T. and Y.C. Ho, (1974), "Informational properties of the Nash solutions of two stochastic nonzero-sum games," Journal of Economic Theory, vol. 7, no. 4, pp. 370-387.

[28] Başar, T. and M. Mintz (1971), "Minimax estimation under generalized quadratic loss", Proc. 12th IEEE Conf. on Decision and Control, Miami, FL, pp. 456-461.

[29] Başar, T. and M. Mintz (1972),"On the existence of linear saddle-point strategies for a two-person zero-sum stochastic game", Proc. 11th IEEE Conf. on Decision and Control, New Orleans, pp. 188-192.

[30] Başar, T. and M. Mintz (1973), "A multi-stage pursuit-evasion game that admits a Gaussian random process as a maximin control policy," Stochastics, vol. 1, no. 1, pp. 25-69.

[31] Başar, T., and G. J. Olsder (1980), "Team-optimal closed-loop Stackelberg strategies in hierarchical control problems," Automatica, vol. 16, no. 4.

[32] Başar, T. and G. J. Olsder (1982), Dynamic Noncooperative Game Theory, Academic Press, London.

[33] Başar, T., and H. Selbuz (1976), "Properties of Nash solutions of a two-stage nonzero-sum game," IEEE Transactions on Automatic Control, vol. AC-21, no. 1, pp. 48-54.

[34] Başar, T. and H. Selbuz (1979a), "A new approach for derivation of closed-loop Stackelberg strategies," in: Proceedings of the 19th IEEE Conference on Decision and Control, San Diego, California, pp. 1113-1118.

[35] Başar, T. and H. Selbuz (1979b), "Closed-loop Stackelberg strategies with applications in the optimal control of multilevel systems," IEEE Transactions on Automatic Control, vol. AC-24, no. 2, pp. 166-179.

[36] Bensoussan, A. (1971), "Saddle points of convex concave functionals," in: H. W. Kuhn, G. P. Szegö, eds., Differential Games and Related Topics, North-Holland, Amsterdam, pp. 177-200.

[37] Bensoussan, A. (1974), "Points de Nash dans le cas de fonctionelles quadratiques et jeux differentiels lineaires a N personnes," SIAM Journal on Control, vol. 12, no. 3, pp. 237-243.

[38] Berkovitz, L. D. (1964), "A variational approach to differential games," in: M. Dresher, L. S. Shapley, A. W. Tucker, eds., Advances in Game Theory, Princeton University Press, Princeton, New Jersey, pp. 127-174.

[39] Bernhard, P. (1979), "Linear quadratic two-person zero sum differential game: Necessary and sufficient condition," Journal of Optimization Theory and Applications, vol. 27, no. 1.

[40] Bertsekas, D. (1976), Dynamic Programming and Stochastic Control, Academic Press, N.Y.

[41] Blackwell, D. and M. A. Girshick (1954), Theory of Games and Statistical Decisions, John Wiley, New York.

[42] Blaquiere, A., F. Gerard and G. Leitmann (1969), Quantitative and Qualitative Games, Academic Press, New York.

[43] Blaquiere, A. (1973), ed., Topics in Differential Games, Nowth-Holland, Amsterdam.

[44] Burger, E. (1966), Einführung in die Theorie der Spiele, Walter de Groyter and Company, Berlin (2.Auflage).

[45] Burrow, J. L. (1978), "A multistage game with incomplete information requiring an infinite memory," Journal of Optimization Theory and Applications, vol. 24, pp. 337-360.

[46] Case, J. H. (1967) Equilibrium Points of N-person Differential Games, Ph.D. Thesis, University of Michigan, Department of Industrial Engineering, TR no. 1967-1, Ann Arbor, Michigan.

[47] Case, J. H. (1969), "Toward a theory of many player differential games", SIAM Journal on Control, vol. 7, no. 2.

[48] Case, J. H. (1971), "Applications of the theory of differential games to economic problems", in: H. W. Kuhn, G. P. Szegö, eds., Differential Games and Related Topics, North-Holland, Amsterdam, pp. 345-371.

[49] Case, J. H. (1979), Economics and the Competitive Process, New York University Press, New York.

[50] Castanon, D. A. (1976), Equilibria in Stochastic Dynamic Games of Stackelberg Type, Ph.D. Thesis, M.I.T., Electronic Systems Laboratory, Cambridge, Massachusetts.

[51] Chen, C. I. and J. B. Cruz, Jr. (1972), "Stackelberg solutions for two-person games with biased information patterns," IEEE Transactions on Automatic Control, vol. AC-17, no. 5, pp. 791-798.

[52] Clemhout, S., G. Leitmann and H.Y. Wan, Jr. (1973), "A differential game model of oligopoly," Journal of Cybernetics, vol. 3, no. 1, pp. 24-39.

[53] Cournot, A. (1838), Recherches sur les principes mathematiques de la theorie des richesses, Hachette, Paris, (English Edition published in 1960 by Kelley, New York, under the title "Researches into the mathematical principles of the theory of wealth," translated by N. T. Bacon.

[54] Cruz, J. B., Jr. (1975), "Survey of Nash and Stackelberg equilibrium strategies in dynamic games," Annals of Economic and Social Measurement, vol. 4, no. 4, pp. 339-344.

[55] Cruz, J. B. Jr. (1978), "Leader-follower strategies for multilevel systems," IEEE Transactions on Automatic Control, vol. AC-23, no. 2, pp. 244-254.

[56] Da Cunha, N.O. and E. Polak (1967), "Constrained minimization under vector-valued criteria in finite dimensional spaces," J. Math. Analysis and Applications, vol. 19, pp. 103-124.

[57] Van Damme, E. E. C. (1980), "History-dependent equilibrium points in dynamic games," Memorandum COSOR 80-05, Dept. of Math., Eindhoven University of Technology, Netherlands.

[58] Fleming, W. H. (1961), "The convergence problem for differential games," Journal of Mathematical Analysis and Applications, vol. 3, pp. 102-116.

[59] Fleming, W. H. (1964), "The convergence problem for differential games II," in: M. Dresher, L. S. Shapley, A. W. Tucker, eds., Advances in Game Theory, Annals of Mathematics Studies, no. 52, Princeton University Press, Princeton, New Jersey, pp. 195-210.

[60] Friedman, A. (1971), Differential Games, John Wiley, New York.

[61] Friedman, A. (1972), "Stochastic differential games," Journal of Differential Equations, vol. 11, no. 1, pp. 79-108.

[62] Friedman, J. W. (1977), Oligopoly and the Theory of Games, North-Holland, Amsterdam.

[63] Grote, J. D. (1975), The Theory and Applications of Differential Games, Reidel Publishing Company, The Netherlands.

[64] Hagedorn, P., H. W. Knobloch and G. J. Olsder, (1977), eds., Differential Games and Applications, Lecture Notes in Control and Information Sciences, vol. 3, Springer-Verlag, Berlin

[65] Haurie, A. and R. Lhousni (1979), "A duopoly model with variable leadership Stackelberg solution," presented at International Atlantic Economic Conference, Salzburg, Austria.

[66] Ho, Y. C. (1969), ed., Proceedings of the First International Conference on the Theory and Applications of Differential Games, Amherst, Massachusetts.

[67] Ho, Y. C. (1970), "Survey paper: Differential games, dynamic optimization and generalized control theory," Journal of Optimization Theory and Applications, vol. 6, no. 3, pp. 179-209.

[68] Ho, Y. C. (1980), "Team decision theory and information structures," Proceeding of the IEEE, vol. 68, no. 6, pp. 644-654.

[69] Ho, Y. C. and K. C. Chu (1972), "Team decision theory and information structures in optimal control problems, Part I," IEEE Transactions on Automatic Control, AC-17, pp. 15-22.

[70] Ho, Y. C. and K. C. Chu (1973), "On the equivalence of information structures in static and dynamic teams," IEEE Transactions on Automatic Control, vol. AC-18, pp. 187-188.

[71] Ho, Y. C. and K. C. Chu, (1974), "Information structures in many person optimization problems," Automatica, vol. 10, pp. 341-351.

[72] Ho, Y. C., A. E. Bryson, Jr. and S. Baron (1965), Differential games and optimal pursuit-evasion strategies," IEEE Transactions on Automatic Control, vol. AC-10, no. 4, pp. 385-389.

[73] Ho, Y. C., M. P. Kastner and E. Wong (1978),"Teams, market signalling and information theory," IEEE Transactions on Automatic Control, vol. AC-23, pp. 305-311.

[74] Ho, Y. C., P. B. Luh and R. Muralidharan (1981), "Information structure, Stackelberg games and incentive controllability, IEEE Transactions on Automatic Control, vol. AC-26, no. 2, pp. 454-960.

[75] Ho, Y. C., P. B. Luh and G. J. Olsder (1980), "A control-theoretic view on incentives," Proceedings of the Fourth International Conference on Analysis and Optimization of Systems, Versailles, France.

[76] Ho, Y. C. and S. K. Mitter (1976), ed., Directions in Large Scale Systems, Plenum Press, New York.

[77] Ho, Y. C. and G. J. Olsder (1980), "Differential games: concepts and applications," in: M. Shubik, ed., Proceedings of Mathematical Models of Conflict and Combat: Uses and Theory.

[78] Isaacs, R. (1969), "Differential games: their scope, nature and future," Journal of Optimization Theory and Applications, vol. 3, no. 5.

[79] Isaacs, R. (1975), Differential Games, Second edition, Kruger Publishing Company, Huntington, New York, (First edition: Wiley, New York, 1965).

[80] Jacobson, D. H. (1977), "On values and strategies for infinite-time linear quadratic games," IEEE Transactions on Automatic Control, vol. AC-22, pp. 490-491.

[81] Krassovski, N. and A. Soubbotine (1977), Jeux Differentiels, Editions Mir, Moscow, (original Russian edition published in 1974).

[82] Kuhn, H. W. and G. P. Szegö (1971), eds., Differential Games and Related Topics, North-Holland, Amsterdam.

[83] Kurtaran, B. Z. (1975), "A concise derivation of the LQG one-step-delay sharing problem solution," IEEE Transations on Automatic Control, vol. AC-20, no. 6, pp. 808-810.

[84] Kydland, F. (1975), "Noncooperative and dominant player solutions in discrete dynamic games," International Economic Review, vol. 16, no. 2, pp. 321-335.

[85] Leitmann, G. (1974), Cooperative and Noncooperative Many Player Differential Games, CISM Monograph no. 190, Springer-Verlag, Vienna.

[86] Leitmann, G. (1976), ed., Multicriteria Decision Making and Differential Games, Plenum Press, New York.

[87] Leitmann, G. (1977), "Many player differential games," in: P. Hagedorn, H. W. Knobloch, G. J. Olsder, eds., Differential Games and Applications, Lecture Notes in Control and Information Sciences, vol. 3, Springer-Verlag, Berlin, pp. 153-171.

[88] Leitmann, G. (1978), "On generalized Stackelberg strategies," Journal of Optimization Theory and Applications, vol. 26, no. 6.

[89] Leitmann, G. and H. Y. Wan, Jr. (1979), "Macro-economic stabilization policy for an uncertain dynamic economy," in: M. Aoki, A. Marzollo, eds., New Trends in Dynamic System Theory and Economics, Academic Press, New York, pp. 105-136.

[90] Levine, J., (1979), "Two-person zero-sum differential games with incomplete information --a Bayesian model," in: P. T Liu, E. O. Roxin, eds., Differential Games and Control Theory III, Proceedings of the 3rd Kingston Conference 1978, Part A. Lecture Notes in Pure and Applied Mathematics, vol. 44, Plenum Press, New York, pp. 119-151.

[91] Levitan, R. E. and M. Shubik (1971), "Noncooperative equilibria and strategy spaces in an oligopolistic market," in: H. W. Kuhn, G. P.Szegö, eds., Differential Games and Related Topics, North-Holland, Amsterdam, pp. 429-448.

[92] Lukes, D. L. (1971), "Equilibrium feedback control in linear games with quadratic costs," SIAM Journal on Control, vol. 9, no. 2, pp. 234-252.

[93] Lukes, D. L. and D. L. Russell (1971), "A global theory for linear quadratic differential games,", Journal of Mathematical Analysis and Applications, vol. 33, pp. 96-123.

[94] Mageirou, E. F. (1976), "Values and strategies for infinite time linear quadratic games," IEEE Transactions on Automatic Control, vol. AC-21, no. 4, pp. 547-550.

[95] Mageirou, E. F. (1977), "Iterative techniques for Riccati game equations," Journal of Optimization Theory and Applications, vol. 22, no. 1, pp.51-61.

[96] Maitra, A. and T. Parthasarathy (1970a), "On stochastic games," Journal of Optimization Theory and Applications, vol. 5, pp. 289-300.

[97] Maitra, A. and T. Parthasarathy (1970b), "On stochastic games, II," Journal of Optimization Theory and Applications, vol. 8, pp. 154-160.

[98] Nash, J. (1951), "Noncooperative games," Annals of Mathematics, vol. 54, no. 2, pp. 286-295.

[99] Von Neumann, J. (1928), "Zur theorie der Gesellschaftspiele," Mathematische Annalen, vol. 100, pp. 295-320.

[100] Von Neumann, J. and O. Morgenstern (1947), Theory of Games and Economic Behavior, Second edition, Princeton University Press, Princeton, New Jersey.

[101] Okuguchi, K. (1976), Expectations and Stability in Oligopoly Models, Lecture Notes in Economics and Mathematical Systems, no. 138, Springer-Verlag, Berlin.

[102] Olsder, G. J. (1976), "Some thoughts about simple advertising models as differential games and the structure of coalitions," in: Y.C. Ho, S. K. Mitter, eds., Direction in Large Scale Systems, Plenum Press, New York, pp. 187-206.

[103] Olsder, G.J. (1977a), "Information structures in differential games," in: E. O. Roxin, P. T. Liu, R. L. Sternberg, eds., Differential Games and Control Theory II, Marcel Dekker, New York, pp. 99-136.

[104] Olsder, G. J. (1977b), "On observation costs and information structures in stochastic differential games," in: P. Hagedorn, H. W. Knobloch, G. J. Olsder, eds., Differential Games and Applications, Lecture Notes in Control and Information Sciences, vol. 3, Springer-Verlag, Berlin, pp. 172-185.

[105] Owen, G. (1968), Game Theory, Saunders, Philadelphia, Pennsylvania.

[106] Papavassilopoulos, G. P. and J. B. Cruz, Jr. (1979a), "Nonclassical control problems and Stackelberg games," IEEE Transactions on Automatic Control, vol. AC-24, no. 2, pp. 155-166.

[107] Papavassilopoulos, G. P. and J. B. Cruz, Jr. (1979b), "On the existence solutions to coupled matrix Riccati differential equations in linear quadratic games," IEEE Transactions on Automatic Control, vol. AC-24, no. 1.

[108] Papavassilopoulos, G. P. and J. B. Cruz, Jr. (1980), "Sufficient conditions for Stackelberg and Nash strategies with memory," Journal of Optimization Theory and Applications, vol. 31, pp. 233-260.

[109] Papavassilopoulos, G. P., J. V. Medanic and J. B. Cruz, Jr. (1979), "On the existence of Nash strategies and solutions to coupled Riccati equations in linear quadratic games," Journal of Optimization Theory and Applications, vol. 28, no. 1, pp. 49-76.

[110] Parthasarathy, T. and T. E. S. Raghavan (1971), Some Topics in Two-person Games, Elsevier, New York.

[111] Parthasarathy, T. and T. E. S. Raghavan (1975), "Existence of saddle-points and Nash equilibrium points for differential games," SIAM Journal on Control, vol. 11, pp. 100-112.

[112] Parthasarathy, T. and M. Stern (1977), "Markov games --a survey," in: E. O. Roxin, P. T. Liu, R. L. Sternberg, eds., Differential Games and Control Theory II, Marcel Dekker, New York, pp. 1-46.

[113] Pindyck, R. S. (1977), "Optimal economic stabilization policies under decentralized control and conflicting objectives," IEEE Transactions on Automatic Control, vol. AC-22, no. 4, pp. 517-530.

[114] Plasmans, J. E. J. (1980), "Interplay: a linked model for economic policy in the EEC," in : Proceedings of the 3rd IFAC/IFORS Conference on Dynamic Modelling and Control of National Economies, Warszawa, Poland.

[115] Pontryagin, L. S. (1967), "Linear differential games, I," Soviet Math. Doklady, vol. 8, pp. 769-771.

[116] Radner, R. (1962), "Team decision problems," Annals of Math. Statistics, vol. 33, pp. 857-881.

[117] Richardson, L. F. (1960), Arms and Insecurity, The Boxwood Press, Chicago.

[118] Rosen, J. B. (1965), "Existence and uniqueness of equilibrium points for concave N-person games," Econometrica, vol. 33, pp. 520-534.

[119] Salman, M. and J. B. Cruz, Jr. (1981), "An incentive model of duopoly with government coordination," Automatica, vol. 17.

[120] Sandell, N. R., Jr., and M. Athans (1974), "Solution of some nonclassical LQG stochastic decision problems," IEEE Transactions on Automatic Control, AC-19, 108-116.

[121] Scalzo, R. C. (1974), "N-person linear quadratic differential games with constraints," SIAM Journal on Control, vol. 12, no. 3, pp. 419-425.

[122] Schmitendorf, W. E. (1970), "Existence of optimal open-loop strategies for a class of differential games," Journal of Optimization Theory and Applications, vol. 5, no. 5.

[123] Shapley, L. S. (1953), "Stochastic Games," Proceedings of the National Academy of Sciences, USA, vol. 39, pp. 1095-1100.

[124] Shiau, T. H. and P. R. Kumar (1979), "Existence and value and randomized strategies in zero-sum discrete-time stochastic dynamic games", Mathematics research report no. 79-17 of the University of Maryland,Baltimore County, Maryland.

[125] Simaan, M. and J. B. Cruz, Jr. (1973a), "On the Stackelberg strategy in nonzero-sum games," Journal of Optimization Theory and Applications, vol. 11, no. 5, pp. 533-555.

[126] Simaan, M. and J. B. Cruz, Jr. (1973b), "Additional aspects of the Stackelberg strategy in nonzero-sum games," Journal of Optimization Theory and Applications, vol. 11, no. 6, pp. 613-626.

[127] Simaan, M. and J. B. Cruz, Jr. (1973c),"Sampled-data Nash controls in nonzero-sum differential games," International Journal of Control, vol. 17, no. 6, pp. 1201-1209.

[128] Simaan, M. and J. B. Cruz, Jr. (1973d), "A Stackelberg solution for games with many players", IEEE Transactions on Automatic Control, vol. AC-18, no. 3, pp. 322-324.

[129] Simaan, M. and J. B. Cruz, Jr. (1975a), "Nash equilibrium strategies for the problem of armament race and control," Management Science, vol. 22, pp. 96-105.

[130] Simaan, M. and J. B. Cruz, Jr. (1975b), "Formulation of Richardson's model of arms race from a differential game viewpoint," The Review of Economic Studies, vol. XLII(1), pp. 67-77.

[131] Simaan, M. and J. B. Cruz, Jr. (1976), "Equilibrium concepts for arms race problems," in Mathematical Systems in International Relations Research, J. V. Gillespie and D. A. Zinnes, editors, Praeger Publishers, Inc., New York, pp. 342-356.

[132] Simaan, M. and T. Takayama (1978), "Game theory applied to dynamic duopoly problems with production constraints," Automatica, vol. 14, pp. 161-166.

[133] Sobel, M. J. (1971), "Noncooperative stochastic games," Annals of Mathematical Statistics, vol. 42, pp. 1930-1935.

[134] Von Stackelberg, H. (1934), Marktform und Gleichgewicht, Springer, Vienna. (An English translation has appeared in 1952 with the title "The Theory of the Market Economy" which was published by Oxford University Press, Oxford).

[135] Starr, A. W. and Y. C. Ho (1969a), "Nonzero-sum differential games," Journal of Optimization Theory and Applications, vol. 3, no. 3, pp. 184-206.

[136] Starr, A. W. and Y. C. Ho (1969b), "Further properties of nonzero-sum differential games," Journal of Optimization Theory and Applications, vol. 3, no. 3, pp. 207-219.

[137] Tolwinski, B. (1978a), "Numerical solutions of N-person nonzero-sum differential games," Control and Cybernetics, vol. 7, no. 1, pp.37-50.

[138] Tolwinski, B. (1978b), "On the existence of Nash equilibrium points for differential games with linear and nonlinear dynamics," Control and Cybernetics, vol. 7, no. 3, pp. 57-69.

[139] Tolwinski, B. (1980), "Stackelberg solution of dynamic game with constraints," in: Proceedings of the 19th IEEE Conference on Decision and Control, Albuquerque, New Mexico.

[140] Tolwinski, B. (1981a), "Closed-loop Stackelberg solution to multistage linear quadratic game," Journal of Optimization Theory and Applications.

[141] Tolwinski, B. (1981b), "Equilibrium solutions for a class of hierarchical games," in: J. Outenbaum, M. Niergodka, eds, Applications of Systems Theory to Economics, Management and Technology, PWN, Warsaw.

[142] Uchida, K. (1978), "On the existence of Nash equilibrium point in N-person nonzero-sum stochastic differential games," SIAM Journal on Control and Optimization, vol. 16, no. 1, pp. 142-149.

[143] Uchida, K. (1979), "A note on the existence of a Nash equilibrium point in stochastic differential games," SIAM Journal on Control and Optimization, vol. 17, no. 1. pp. 1-4.

[144] Varaiya, P. P. (1967), "The existence of solutions to a differential game," SIAM Journal on Control, vol. 5, pp. 153-162.

[145] Varaiya, P. P. (1970), "N-person nonzero-sum differential games with linear dynamics," SIAM Journal on Control, vol. 8, pp. 441-449.

[146] Varaiya, P. P. (1976), "N-player stochastic differential games," SIAM Journal on Control and Optimization, vol. 14, pp. 538-545.

[147] Varaiya, P. P. and J. Walrand (1978), "On delayed sharing patterns," IEEE Transactions on Automatic Control, vol. AC-23, pp. 443-445.

[148] Vidyasagar, M. (1976), "A new approach to N-person, nonzero-sum, linear differential games," Journal of Optimization Theory and Applications, vol. 18, pp. 171-175.

[149] Vorob'ev, N. H. (1977), Game Theory, Springer-Verlag, Berlin.

[150] Wilson, D. J. (1977), "Differential games with no information," SIAM Journal on Control and Optimization, vol. 15, pp. 233-246.

[151] Witsenhausen, H. S. (1968), "A counter example in stochastic optimum control," SIAM Journal on Control, vol. 6, no. 1, pp. 138-147.

[152] Witsenhausen, H. S. (1971a), "On information structures, feedback and causality," SIAM Journal on Control, vol. 9, pp. 149-160.

[153] Witsenhausen, H. S. (1971b), "On the relations between the values of a game and its information structure," Information and Control, vol. 19, no. 3.

[154] Witsenhausen, H. S. (1971c), "Separation of estimation and control for discrete-time systems," Proc. IEEE, vol. 59, pp. 1557-1566.

[155] Witsenhausen, H. S. (1973), "A standard form for sequential stochastic control," Math. Systems Theory, vol. 7, pp. 5-11.

[156] Witsenhausen, H. S. (1975), "The intrinsic model for discrete stohastic control: some open problems," in: A. Bensoussan and J. L. Lions, eds, Lecture Notes in Economics and Mathematical Systems, vol. 107, pp. 322-335.

[157] Willman, W. W. (1969), "Formal solutions for a class of stochastic pursuit-evasion games," IEEE Transactions on Automatic Control, vol. AC-14, no. 5, pp. 504-509.

[158] Wishart, D. M. G. and G. J. Olsder, (1979), "Discontinuous Stackelberg Solutions," International Journal of System Sciences, vol. 10, no. 12, pp. 1359-1368.

[159] Yoshikawa, T. (1975), "Dynamic programming approach to decentralized control problems," IEEE Transactions on Automatic Control, vol. AC-20, pp. 796-797.

[160] Yoshikawa, T. and H. Kobayashi (1978), "Separation of estimation and control for decentralized stochastic control systems," Automatica, vol. 14, no. 6, pp. 623-628.

[161] Zheng, Y. P. and T. Başar (1981), "Existence and derivation of optimal affine incentive schemes for Stackelberg games with partial information: A geometric approach," submitted for publication.

policy is changed, the reactions of the subsumed decision makers may change so that the fixed model being used may not be satisfactory anymore. It would be preferable to explicitly model the presence of the other decision makers.

For situations where cooperation among decision-makers is desirable, the concept of Pareto optimality is appropriate. However, in non-cooperative situations the Nash equilibrium concept is more natural. Hierarchies in decision-making lead to the concept of Stachelberg or leader-follower strategies. These concepts are described in this chapter for both deterministic and stochastic systems.

A critical consideration in multi-person optimization problems is the information structure. In contrast to single person decision making which necessarily involves centralized information, the multi-person decision-making problem may involve decentralized information. Furthermore, the

END